CHAPTER 2

FRPAs and High-Gain Directional Antennas

Basrur Rama Rao

2.1 Categories of GNSS Antennas

Many types of GNSS antennas have been developed in recent years to make them suitable for different applications. Since their designs and performance requirements vary depending on their application, they have been grouped into six different categories in this book:

1. FRPA;
2. High-gain directional antennas;
3. GPS adaptive antennas;
4. Multiband antennas;
5. Handset antennas;
6. Active antennas.

The design and performance of each type of antenna will be described here and throughout the rest of the book.

This chapter discusses the design of the first two varieties: FRPAs and high-gain directional antennas. Figure 2.1 shows a representative sample of these antennas. FRPAs are the most popular and widely used of all GNSS antennas. There are many different types of these antennas; hence they will need two chapters—this chapter and Chapter 3—to fully cover all the important types. This chapter will discuss the following FRPA designs: microstrip patch antennas, which are the most ubiquitous of all GNSS antennas, and quadrifilar helix antennas (QHAs), which are also popular for handheld receivers and crossed “drooping” bow-type dipoles, which provide relatively wider bandwidths and good circular polarization. Spiral antennas such as the Archimedean spiral antenna will be discussed in Chapter 3 under multiband antennas since they are ultrawideband and can cover the entire GNSS band from 1.1 to 1.6 GHz. The conical spiral antenna is also a high-gain di-
Directional antennas have radiation characteristics that are distinctly different from FRPAs. FRPAs are receiving antennas that have a broad antenna pattern for acquiring four or more satellites needed for a PVT solution. Directional antennas can be used as either receiving or transmitting antennas. When operated in the receiving mode they have high gain but narrowbeam patterns for selecting only signals of interest (SOI)—the GNSS satellite signals—while rejecting signals not of interest (SNOI) such as multipath signals. When used as transmitting antennas they generate a high-gain, narrowbeam pattern that can be pointed towards either a specific target or region of operation. Three types of GNSS directional antennas are described in this chapter: (1) helical antennas, (2) reflector antennas, and (3) beamforming antenna arrays. Helical antennas are used as elements in transmitting antenna arrays in GNSS satellites for generating an Earth-coverage beam [1–3] and for other special GNSS applications such as transmit antennas for pseudolites [4] and also during laboratory testing. Several large reflector antennas, which range in diameter from 1.8 to 110m, have also been used as receiving antennas for monitoring GNSS signals transmitted by recently launched satellites (such as GIOVE-A&B), and the Compass M1 [5–8]. Beamforming antenna arrays are another type of directional receiving antenna capable of steering four or more beams towards selected satellites, thereby greatly increasing the signal-to-noise ratio while simultaneously reducing unwanted signals such as multipath [9–12].

2.1.1 FRPA

An FRPA antenna has a nearly omnidirectional pattern in the upper hemisphere and is designed for acquiring almost all, but at least a minimum of four, satellites visible...
to the antenna above a certain masking angle. A representative elevation plane pattern of a conventional FRPA antenna of the first category is shown in Figure 2.2. It is expected to meet a number of desired performance requirements with the objective of being able to provide high precision in GPS measurement. These requirements are described in several recent publications [13, 14] and can change depending on the intended application. The antenna is expected to be RHCP so as to efficiently receive signals from GNSS satellites. The antenna is also expected to provide a better than the minimum required gain over much of the upper hemisphere covering 360° in azimuth and from zenith down to a masking angle of generally between 5° or 10° in elevation; this assures that the receiver has high satellite availability and is able to acquire at least four or more satellites within its view. Good PDOPs, well below the required maximum limit of six, can be achieved if the antenna is capable of receiving signals from low-elevation satellites that are also widely separated in azimuth. For airborne GNSS antennas the gain requirements at low-elevation angles are particularly daunting [20] since they require −3 to −4.5 dBic of RHCP gain at 10° and 5° in elevation, respectively. Achieving good gain and good circular polarization (CP) axial ratio at such low-elevation angles is generally a challenge since in most types of GNSS antennas that are located on metal ground planes (or the aircraft fuselage for avionics antennas) the gain drops off sharply from its peak value at zenith as the elevation decreases. The metallic ground plane nulls out the horizontally polarized component at its surface (i.e., the horizon); hence, the polarization of the antenna becomes linear instead of RHCP and is oriented vertical to the ground plane representing a further 3-dB loss in gain. These antennas should also preferably be able to discriminate against multipath by having good front-to-back ratios with reduced gain at low-elevation angles at and below the horizon. Since multipath reflections change the polarization state from RHCP to LHCP, these antennas are required to have good axial ratios with low-LHCP cross-polarization level at low-elevation angles where multipath effects are most prominent. The desired axial ratio expected from airborne GNSS antennas should be no

Figure 2.2  Representative measured elevation plane pattern of an FRPA for GNSS.
greater than 3 dB for all operating frequencies at elevation angles greater than 10° nor exceed 6 dB for all operating frequencies for elevations between 5° and 10° [13, 14]. The antenna therefore needs to have a cross-polarization ratio of 15.3 dB to be able to meet the 3 dB in the axial ratio specification; the cross-polarization ratio is defined by the ratio of the power density in the cross-polarized LHCP component of the incoming signal to that of the principal RHCP component. High-precision geodetic quality antennas when used for carrier phase tracking are also required to have a stable phase center that varies only minimally with elevation. This same requirement is also needed for GPS-based attitude determination systems. The antenna is also required to limit the group delay variation with frequency over the operating bandwidth of the antennas. Group delay becomes particularly important in maintain the fidelity of BOC codes that are currently being used for M code in modernized GPS as well as for several of the frequency bands in Galileo [15].

2.2 Microstrip Antennas

Microstrip antennas, commonly called patch antennas, are the most popular type of GNSS antennas used in a large variety of civilian and military systems. Their very low-profile, compact size, ability to conform their shape to that of the host surface, ease of obtaining RHCP, and low cost of manufacture gives them unique advantages that are difficult to match for GNSS applications with any other antenna design. They are universally used in avionics since their low profile and compactness easily allows them to meet the ARINC 743 size requirements, which is a voluntary aircraft industry specification; this requirement restricts the lateral sizes of the antenna to be no more than 4.7” × 3” and their height to no more than 0.73”. The ARINC 743 cross section is shown in Figure 2.3. Another form specification is called “teardrop,” which is even smaller than ARINC 743. These same qualities also makes them best suited for use in all types of adaptive antenna arrays used in GPS military navigation systems for combating jamming and interference and for beamforming. Patch antennas are also popular for handset antennas since miniaturized antennas can be built using high-dielectric constant ceramic substrates; this allows them to be easily integrated into a variety of popular handheld navigation devices such as cell phones and PDAs. Shorted annular ring microstrip antennas have been proposed for multipath limitation; these are smaller in size, less complex to manufacture, and lower cost than the more expensive choke ring antennas. Microstrip antennas will therefore be discussed here in much greater detail than the other GNSS antennas considered in this book owing to their importance and popularity.

The basic microstrip antenna consists of a metallic conductor of a specific shape that is etched on the top surface of dielectric substrate that is physically bigger than the metallic patch. Both the metallic patch and the substrates are placed over an even larger metallic ground plane that can sometimes be many times the GNSS wavelength. The copper cladding of the bottom surface of the dielectric substrate becomes the central part of the ground plane. Single-band microstrip antennas use a single dielectric substrate but dual- and triple-band patch antennas can use two or more substrates.

Figure 2.4(a) shows a picture of a basic RHCP microstrip patch antenna; the conducting ground plane underneath the patch antenna is not shown in this figure.
RHCP is achieved by using two coaxial probes that are clearly visible in this picture. A schematic sketch of this antenna is shown in Figure 2.4(b). The metallic patch and the ground plane are assumed to be good electrical conductors and form the
top and bottom surfaces of a high Q resonant RF cavity that is tuned to resonance at the desired GNSS frequency. The four edges of the patch act as perfect magnetic conductors and form the sides of the cavity. The energy stored inside the resonant cavity leaks out from the edges of the metallic patch. The resonance frequency of the patch antenna for a desired cavity mode is dependent on the size and shape of the conducting patch, the dielectric constant $\varepsilon_r$ of the dielectric substrate, and the thickness of the substrate layer. The directional properties of radiation pattern and its polarization are determined by the electromagnetic fields of the cavity modes generated within the RF cavity and the method of feeding the patch antenna for receiving the satellite signals. To be suitable for GNSS applications the patch antenna needs to have all principal characteristics of a typical FRPA antenna noted earlier.

The three most commonly used shapes for the metallic patches to achieve maximum circular symmetry in azimuth are a square, a circle, or an annular ring, as shown in Figures 2.5(a), (b), and (c), respectively.

A majority of GPS microstrip antennas currently used in low-cost, commercial GPS receivers operate only over a single-frequency band—the GPS L_1 band with a center frequency of 1.5754 GHz—and use a single layer of the dielectric substrate. Some of these antennas have a 2-MHz bandwidth only just sufficient to receive the C/A code. These single-band antennas are the simplest to design and are the least expensive. Dual-band microstrip antennas that operate at either the L_1 and L_2 bands or the L_1 and L_5 bands and also triple-band antennas that operate in all three frequency bands—L_1, L_2, and L_5—will be needed soon to meet the demands of the modernized GPS. Multiband patch antennas designed to meet these requirements are discussed later in this chapter and in Chapter 3. These multiband patch antennas consist of a combination of two or more patch antennas with each patch antenna resonating at a different frequency, generally either stacked on top of each other or parasitically coupled to one another. These antennas may require a bandwidth of 20 MHz in each band or 24 MHz if reception of a GPS M code signals is needed. All the microstrip antennas shown in Figure 2.5(a), (b), and (c) are dual-band “stacked patch” antennas, whose design will be explained in greater detail later; they operate in the GPS L_1 and L_2 bands and use two dielectric substrate layers—one for each patch antenna.

![Figure 2.5](image-url)  (a) Square, (b) circular, and (c) annular ring GPS microstrip antennas.
2.2.1 Selection of the Dielectric Substrate for Microstrip Antennas

The selection of a suitable dielectric substrate is a critical first step in the design of microstrip patch antennas and involves a trade-off between different application-driven requirements such as size and height, bandwidth, and gain coverage in the upper hemisphere. Small-size antenna elements with a low profile are required in avionics and for miniaturized adaptive antenna arrays used in military airborne navigation systems; very compact antennas that can be unobtrusively integrated into the receiver are used in handsets. A substrate with a higher dielectric constant reduces the size of the patch, but at the cost of decreasing both the bandwidth and gain of the antenna with increased radiation near the horizon from surface waves.

A variety of substrates with dielectric constants ranging from as low as 1.07 to as high as 88 are available for building GNSS antennas depending on the application. More popular substrate materials are listed in Table 2.1 along with their dielectric constant and manufacturer. Three of these substrates with different dielectric constants will be selected later to illustrate the effects of the dielectric properties of the substrate on the performance of the GNSS patch antenna. Many of these substrates have electrodeposited copper cladding with a thickness of \( \frac{1}{2} \) mil to 1 mil on their upper and lower surfaces; this corresponds to a copper foil thickness of 0.0007 inches to 0.0014 inches, respectively. The desired shape and size of the patch can be obtained by either photo etching or milling the top copper cladding and the bottom copper cladding serves as the ground plane. High dielectric constant ceramic substrates with a single-probe feed are particularly popular for compact, narrowband GNSS antennas used in handsets. Since these antennas have very narrow bandwidths, the temperature stability of the dielectric constant is very important in ceramic substrates for preventing detuning if a large variation in the ambient temperature were to occur, as for example in avionics systems. Lower dielectric constant substrates such as foam or foam derivatives are also often used in conjunction with other higher dielectric constant substrates for improving bandwidth and are also used as substrates for multiband antennas and in antenna feeds. More

Table 2.1 Dielectric Substrates for GNSS Microstrip Antennas

<table>
<thead>
<tr>
<th>Name of Substrate</th>
<th>Dielectric Constant</th>
<th>Loss Tangent</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rohacell Foam</td>
<td>1.07</td>
<td>0.001</td>
<td>Rohm</td>
</tr>
<tr>
<td>Arlon Foam-Clad 100</td>
<td>1.15–1.35</td>
<td>0.002–0.004</td>
<td>Arlon; <a href="http://www.arlom-med.com">www.arlom-med.com</a></td>
</tr>
<tr>
<td>Duroid 5870</td>
<td>2.35</td>
<td>0.005</td>
<td>Rogers Corp.; <a href="http://www.rogerscorporation.com">www.rogerscorporation.com</a></td>
</tr>
<tr>
<td>RO3003</td>
<td>3.00</td>
<td>0.0013</td>
<td>Rogers Corp.</td>
</tr>
<tr>
<td>TMM4</td>
<td>4.50</td>
<td>0.0017</td>
<td>Rogers Corp.</td>
</tr>
<tr>
<td>TMM6</td>
<td>6.0</td>
<td>0.0018</td>
<td>Rogers Corp.</td>
</tr>
<tr>
<td>TMM10</td>
<td>9.2</td>
<td>0.0017</td>
<td>Rogers Corp.</td>
</tr>
<tr>
<td>RO3010</td>
<td>10.2</td>
<td>0.0023</td>
<td>Rogers Corp.</td>
</tr>
<tr>
<td>TMM13i</td>
<td>12.78</td>
<td>0.002</td>
<td>Rogers Corp.</td>
</tr>
<tr>
<td>SM200 (ceramic)</td>
<td>20</td>
<td>0.001</td>
<td>Kyocera North America; americas.kyocera.com</td>
</tr>
<tr>
<td>SB 350 (ceramic)</td>
<td>35</td>
<td>0.001</td>
<td>Kyocera North America</td>
</tr>
<tr>
<td>D88 (ceramic)</td>
<td>88 ± 2</td>
<td>Temperature coefficient = 0 ± 5 ppm(^\circ) C(^2)</td>
<td>Morgan Electro Ceramics, U.K.; <a href="http://www.morgan-electroceramics.com">www.morgan-electroceramics.com</a></td>
</tr>
</tbody>
</table>
recently textile materials, including synthetic fabrics also called e-textiles, have been considered as flexible substrates for building body-worn, circularly polarized GPS antennas [17].

### 2.2.2 Effects of Surface Waves on Microstrip GNSS Antennas

All microstrip antennas need a grounded dielectric substrate as indicated in Figure 2.4; they therefore generate surface waves [24], which have a significant impact on performance when used in GNSS. As implied by its name these are electromagnetic modes of propagation that are trapped in the grounded dielectric substrate and travel via successive reflections between the dielectric-air boundary and the metallic ground plane underneath the substrate, as shown in Figure 2.6 Their field amplitudes decrease slowly with distance $r$, the propagation distance. Surface waves propagate until they reach the truncated edge of the dielectric substrate or the edge of a finite (substrate-covered) ground plane where they are either diffracted or reflected. The diffracted signals then interact with the primary space wave radiation from the patch antenna and can cause numerous undesirable effects on GNSS performance. These effects include increased multipath from antenna back lobes, higher cross-polarization levels, phase variations, and ripples in the antenna pattern that are noticeable even at higher-elevation angles close to zenith. Surface waves have a deleterious impact on the performance of a GNSS adaptive antenna array due to increased mutual coupling between adjacent elements in the array; these effects are illustrated in Figure 2.6 and are also discussed in greater detail in Chapter 5. In beamforming arrays they can cause a change in the antenna pattern response of each array element requiring careful calibration of the amplitude and phase of each array element as a function of azimuth and elevation angle to compensate for their effects on beamforming. They can also produce enhanced interaction with the human operator in handset antennas.

Surface waves in microstrip antennas can be divided into two types: transverse magnetic (TM) mode and the transverse electric mode (TE) mode [18]. In TM surface waves, the magnetic field is parallel to the surface, whereas the electric field forms loops that extend vertically out of the surface. In TE surface waves the

![Figure 2.6](attachment:image.png)
2.2 Microstrip Antennas

The electric field is parallel to the surface and the magnetic field forms vertical loops out of the surface. Cutoff frequencies for the different surface wave modes are given by:

\[ f_c = \frac{n c}{4 h \sqrt{\varepsilon_r - 1}} \]

where \( h \) is the thickness of the substrate, \( c \) is the velocity of light = \( 3 \times 10^{10} \) cms per second, and \( \varepsilon_r = \) permittivity of the substrate; \( n = 1, 3, 5 \) for the TE\(_n\) modes and \( n = 0, 2, 4 \) for the TM\(_n\) modes. The lowest order TM\(_0\) mode has no cutoff frequency, can be excited at any frequency, and has the greatest effect on GNSS antennas. It can only be mitigated by using a specially designed shorted annular ring microstrip antenna [19, 20]; this will be discussed later in this chapter under GPS multipath limiting microstrip antennas. Propagation of the lowest order TE\(_1\) mode can however be avoided by keeping the thickness of the dielectric substrate less than \( h_c \) where:

\[ h_c = \frac{0.3 c}{2 \pi f_U \sqrt{\varepsilon_r}} \]

When \( f_U = 1.5754 \) GHz, \( h_c \leq \frac{0.3576}{\sqrt{\varepsilon_r}} \) inches. Surface wave energy propagated by the TE\(_1\) mode can be decreased by keeping the thickness of the substrate below \( h_c \). However, this can reduce the bandwidth and also the radiation efficiency. Alternatively, surface waves can be reduced by selecting a substrate with a lower dielectric constant but this can increase the size of the patch antenna, which is not desirable for many GNSS applications.

Despite its many disadvantages, not all of the surface wave effects created by the grounded dielectric substrate are necessarily detrimental for GNSS use. In an isolated antenna, surface waves can provide some unexpected benefits for GNSS applications such as increasing the gain of the antenna closer to the horizon, thereby promoting the acquisition of low-elevation satellites for improving PDOP. This is generally found difficult to achieve with other low-profile antenna designs where, due to their limited vertical height, they mainly produce broadside radiation directed towards zenith and away from low elevations.

The impact of surface waves on antenna performance is determined by the surface wave efficiency defined by:

\[ \eta_{surf} = 1 - \frac{P_{SUR}}{P_r} \]

where \( P_{SUR} \) is power that is trapped within the surface wave and \( P_r \) is the total radiated power from the antenna. Surface wave efficiency will play a prominent role in the performance parameters of GNSS antennas as discussed later.

2.2.3 Design of Dual-Probe-Fed RHCP Single-Band Microstrip GNSS Antenna

Several types of feeding techniques have been devised for microstrip antennas with the goal of generating RHCP radiation for their use in GNSS systems. Some of these are quite complex as discussed later and can have a profound impact on its performance. We will first consider the design and performance of one of the simplest of the many methods for generating RHCP; namely, two direct-contact coaxial probes feeding a square-shaped RHCP microstrip antenna for operation over a single GNSS frequency band. We will also conduct a parametric study of this antenna to illustrate how various parameters affect its performance when used in GNSS. A schematic diagram of this antenna is shown in Figure 2.4(a). The antenna dimensions of the patch antenna are \( L \) along the \( x \) axis, \( W \) along the \( y \) axis, and \( h \) is the thickness of the dielectric substrate. The patch antenna is placed on top of a
large square metallic ground plane with dimension \( b \). A square patch is needed for generating RHCP, so \( W = L = a \). However, although the dimensions of this square patch along the \( y \) and \( x \) axes are the same, it is nevertheless useful to distinguish between these two sides of the patch by using letters \( W \) and \( L \) to aid in the discussion of TM and TE modal electric fields and their currents. An RHCP-radiated field is obtained by placing the two probes at orthogonal positions within the patch, as shown in Figure 2.4. Each probe is just an extension of the center conductor of the coaxial feed line and is soldered to the conducting patch antenna with the outer conductor of the coaxial line soldered to the ground plane. The coordinates of Probe 1 is \( X = X_p, Y = +a/2 \); and of Probe 2 is \( x = a/2, y = y_p \). \( x_p \) is the distance of the center of Probe 1 from the upper horizontal edge of the patch antenna and \( Y_p \) is the distance of the center of Probe 2 from the left vertical edge of the patch. The locations of these two probes are precisely selected so that the input resistance of each probe is close to 50 ohms and the input reactance is close to zero, ensuring a good impedance match and low return loss to the feeding coaxial cables connected to the GNSS receiver. The two probes are connected to a 90° hybrid such as a branch-line coupler or a coaxial hybrid to generate RHCP signals. Probe 1 has 0° phase, whereas Probe 2 has \(-90°\) phase. The design of the branch-line coupler used with a microstrip antenna will be discussed later in this chapter. The electric fields inside the cavity are directed along the \( z \) axis and are independent of the \( z \) coordinate but vary along the \( x \) and \( y \) axes depending on the electromagnetic modes generated within the resonant microwave cavity of the patch antenna. These can be described in terms of the TM\(_{mn}\) resonant cavity modes identified by their specific double index \((m,n)\). The integer mode index \( m \) of the TM\(_{mn}\) mode is related to the half-cycle variations of the electric field under the square patch over the width \( W \) along the \( x \) axis. The mode index \( n \) is related to the number of half-cycle electric field variations over the length \( L \) parallel to the \( y \) axis. \( L \) is approximately one-half wavelength in the dielectric substrate and the length of the patch \( a \sim \lambda_0/2 \sqrt{\varepsilon_r} \), where \( \lambda_0 \) is the wavelength at the resonant frequency of the antenna and \( \varepsilon_r \) is the relative dielectric constant of the substrate. For the TM\(_{mn}\) cavity mode of the rectangular patch, the \( z \)-directed electric field has the form

\[
E_z(x,y) = A_{mn} \cos \left( \frac{m \pi x}{W_{\text{eff}}} \right) \cos \left( \frac{n \pi y}{L_{\text{eff}}} \right)
\]  

(2.1)

where \( A_{mn} \) is the modal amplitude of the cavity mode \((m, n)\). The walls of the cavity are slightly larger electrically than they are physically due to the fringing electric field at the edges of the patch. To compensate for this the boundaries of the metallic patch are extended outwards and the new dimensions become \( W_{\text{eff}} \) and \( L_{\text{eff}} \) as indicated in (2.1). For GNSS applications, the lowest fundamental order modes, the TM\(_{10}\) or the TM\(_{01}\), are selected to obtain an azimuthally symmetric RHCP pattern. To generate RHCP required for receiving the GNSS signals, the phase of Probe 2 is 0° and the phase of Probe 1 is \(-90°\). Probe 1 excites the fundamental TM\(_{10}\) mode and is polarized linearly along the +\( X \) axis. The electric field for radiation Probe 2, which is resonant in the TM\(_{01}\) mode, is polarized along the +\( Y \) axis. The direction of the magnetic current flows and the radiating fields from each of these two probes is illustrated and explained in greater detail in Figure 5.1. The radiation fields of
this patch antenna from these two pairs of slots parallel to the X and Y axes, respectively, generate orthogonally polarized electric field, which when combined in phase quadrature such as a branch-line coupler generate the required RHCP radiated signals for GNSS applications.

2.2.3.1 Computer Codes for Designing Microstrip Antennas

Despite its structural simplicity, an accurate design of a microstrip antenna that is needed for meeting various GNSS requirements requires the use of computational electromagnetic codes. A discussion of these codes is beyond the scope of this book, but they are well described in handbooks [21, 22] and also in two excellent textbooks devoted exclusively to microstrip antennas [18, 23]. A large number of computer codes are now available to the antenna engineer for designing microstrip antennas; some of the more frequently used codes are listed in Table 2.2. A comparison of the accuracy of some of these codes for microstrip antenna design has been conducted by Pozar et al. [24]. Due to the long computer run times that may be needed, these codes should be considered as suitable for only verifying and optimizing the final design; the starting initial design is often based on simplified computer-aided design (CAD) formulas or on engineering intuition of the antenna designer. Several hardware iterations may also be needed for achieving an optimum design that meets most of the antenna requirements since the computer models may not account for all the nuances in the actual design. Some specific examples of GNSS dual-band microstrip patch antennas designed using the High Fidelity System Simulation (HFSS) code will be described later in this section; the performance of these antennas have been verified through measurements and agree fairly well with design predictions.

2.2.4 A Parametric Study of a Single-Band RHCP Square-Shaped Microstrip Antenna

We will first consider the procedure for designing a basic GNSS microstrip antenna: an RHCP, single-band GNSS microstrip antenna fed with a single substrate layer and two feed probes as shown in Figures 2.4(a) and (b). Designing this simple antenna will help the reader to obtain a better understanding of how the various parameters can first be selected for the initial design needed for meeting the many

<table>
<thead>
<tr>
<th>Table 2.2</th>
<th>Electromagnetic Codes Used for the Design and Analysis of Microstrip Antennas</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name of Software Code</strong></td>
<td><strong>Analysis Technique</strong></td>
</tr>
<tr>
<td>Ensemble (Designer)</td>
<td>Method of moments</td>
</tr>
<tr>
<td>HFSS</td>
<td>Finite element</td>
</tr>
<tr>
<td>IE3D</td>
<td>Method of moments</td>
</tr>
<tr>
<td>Microwave Studio Suite &amp; Microstrips</td>
<td>FDTD TLM</td>
</tr>
</tbody>
</table>

FDTD = finite difference time domain; TLM = transmission line modeling.
important GNSS requirements such as size, location of feed probes in the antenna, bandwidth, radiation efficiency, gain, beam width, polarization axial ratio, and stability of the antenna phase center.

Simple analytical formulas [21] and numerical calculations obtained from a CAD design tool by Sainati [25] will be used to provide a preliminary insight into selecting the parameters. The CAD design tool by Sainati is based on curve fitting of rigorous full-wave solutions and provides good initial estimates that compare well with several actual verification measurements. Three temperature-stable, low-loss dielectric substrates obtained from Table 2.1 will be considered in this study to show their effects on antenna performance; all three are made by the Rogers Corporation and have dielectric constants of 2.35 (Duroid 5870), 6.0 (TMM6) and 9.2 (TMM10). Although slightly less accurate, this approach will provide the reader with good physical insight into the role played by the various design variables in determining the performance of the microstrip antenna. It also serves as an initial starting point in the antenna design whose accuracy can be refined by using more advanced electromagnetic codes mentioned in Table 2.2 or through iterative measurements on prototype antenna models for arriving at a final optimized design. Later in this section we will consider dual-band stacked microstrip antennas designed using the more accurate HFSS code and provide examples of the measured performance of these antennas.

2.2.4.1 Resonance Frequency and Size of a Single-Layer Square-Shaped GNSS Microstrip Antenna

The resonance for the TM$_{10}$ (or the TM$_{01}$) mode is given by

$$f_0 = \frac{c}{2a_{\text{eff}} \sqrt{\varepsilon_r}}$$  \hspace{1cm} (2.2)

where $c$ is the speed of light and $\varepsilon_r$ is the relative permittivity of the dielectric substrate. To account for the fringing of the electric fields of the cavity at the edges of the patch, an effective length $a_{\text{eff}}$ for the dimensions of the square patch while calculating the resonance frequency is chosen as

$$a_{\text{eff}} = a + 2\Delta a$$  \hspace{1cm} (2.3)

The Hammerstad formula for the effective length extension [26] and the effective dielectric constant caused by the fringing fields proposed by Schneider [27] can be used for resonance frequency calculation

$$\Delta a = 0.412 \left[ \frac{(\varepsilon_{\text{eff}} + 0.3)(\frac{a}{b} + 0.264)}{(\varepsilon_{\text{eff}} - 0.258)(\frac{a}{b} + 0.8)} \right]$$  \hspace{1cm} (2.4)

where
Figure 2.7 shows the variation in the size of a square-shaped patch antenna as a function of the thickness of the three selected dielectric substrates with permittivity \( \varepsilon_r = 2.35 \) (Rogers Duroid 5870), \( \varepsilon_r = 6.0 \) (Rogers TMM6), and \( \varepsilon_r = 9.2 \) (Rogers TMM10). Notice that the length of the sides of this square patch is reduced by nearly 50% as the dielectric constant is increased from 2.35 to 9.2. The reduction in size caused by increasing the dielectric constant of the substrate increases the beamwidth but reduces both the gain and bandwidth, as discussed later.

### 2.2.4.2 Locations of Feed Probes in an RHCP Patch Antenna

The location of the two feed probes to generate RHCP is the next critical step in the design of a GNSS patch antenna. As explained earlier, the two feed probes need to be located orthogonally at the correct distances from the edges of vertical and horizontal sides of the patch to obtain a good impedance match to the GNSS receiver. The two probes are connected to a quadrature hybrid, such as branch-line coupler, so as to have equal amplitudes but with a phase of 90 degrees relative to each other to generate RHCP radiation from the patch antenna.

The resonant radiation resistance \( R_r \) of a square or rectangular shape microstrip fed at an edge of the patch is given by [21, pp. 277–279]:

\[
R_r = \frac{V_0^2}{2P_r} = \varepsilon_r \frac{Z_0^2}{120f_2}
\]  

(2.6)
$V_0$ is the voltage across the representative edge slot in the radiating microstrip antenna, $P_r$ is power radiated by the patch, and $Z_0$ is the characteristic impedance of the microstrip line of which the patch is a segment. The power $P_r$ radiated by the antenna can be obtained by integrating the real part of the Poynting vector over the hemisphere above the disk.

$$P_r = \frac{1}{2\eta_0} \int_0^{\frac{2\pi}{2}} \int_0^{\frac{\pi}{2}} \left( |E_\theta|^2 + |E_\phi|^2 \right) r^2 \sin \theta \, d\theta \, d\phi$$

where $E_\theta$ and $E_\phi$ are the radiated electric field components along the $\hat{\theta}$ and $\hat{\phi}$ vector directions of the far field. $I_2$ is a complicated function of several parameters such as $\varepsilon_{re}$ the real part of the complex dielectric constant of the dielectric substrate, $b$ is the thickness of the dielectric substrate, and $W = a$ is the width of the patch antenna [21, pp. 279].

As shown in Figure 2.4(b), the patch antenna is fed with the feed Probe 1 located at a distance on the X axis and at $Y = a/2$ from the top, horizontal radiating edge of the patch; its input resistance is obtained as

$$R_{in} = R_r \cos^2 \left( \frac{\pi x_p}{L} \right)$$

A good impedance match to the receiver is obtained when input resistance of the Probe $R_{in} = 50$ ohms. Similarly the location of the other orthogonal feed Probe 2 is located at a distance $y_p$ along the Y axis and at $X = a/2$ from the vertical, left edge of the patch. $y_p$ can be calculated from (2.8) by substituting $y_p$ for $x_p$ and $L$ for $W$, noting that for a square patch since $L = W = a$. In Figure 2.8 the offset distance of the probe from the edge of the patch, which is $X_p$ for Probe 1 or $Y_p$ for Probe 2, are plotted for three different substrates with dielectric constants $\varepsilon_r = 2.35$, $\varepsilon_r = 6.0$, and $\varepsilon_r = 9.2$ as a function of the thickness of the dielectric substrate. Not surprisingly, the offset distance of the probe from the edge decreases as the dielectric constant of the substrate is increased since the size of the patch antenna also decreases. For a specific substrate the probe offset distance increases as the thickness is increased to about 0.25 cms, but then plateaus out with little change as the thickness is increased further.

### 2.2.4.3 Bandwidth of a GNSS Microstrip Antenna

The bandwidth of the antenna is the frequency range over which a selected performance parameter of the antenna has satisfactory operation. Depending on the specific performance parameter selected, three different definitions of bandwidth are possible: the impedance bandwidth (or return loss bandwidth), gain bandwidth, and axial ratio bandwidth. The impedance or the return loss bandwidth denotes a frequency range over which 89% of the signal power received by the patch antenna is transferred to the GNSS receiver, representing a return loss of $-9.5$ dB (11% reflected power). The gain bandwidth is defined in terms of the frequency range over which the antenna provides a gain better than the minimum threshold gain needed...
to acquire the GNSS satellites within a viewing region covered by the entire upper hemisphere down to the minimum specified low-elevation masking angle. The axial ratio bandwidth is defined by the frequency range in which the antenna is able to meet the maximum cross-polarization axial ratio level. Generally the impedance bandwidth of patch antennas is much narrower than either its gain or axial ratio bandwidth. The return loss bandwidth, which can often be smaller than the gain bandwidth by nearly a factor of 10 [28, pp. 17–18] is therefore a more difficult requirement to meet. The impedance bandwidth is also relatively easier to measure using a vector network analyzer whereas the measurement of gain or axial ratio needs some form of antenna range that is more expensive and less readily available. Hence the bandwidth of GNSS antennas is often defined only in terms of impedance bandwidth, although in terms of actual GNSS performance the gain bandwidth of the antenna is probably a far more meaningful parameter for determining satellite acquisition and also the signal quality (carrier-to-power noise ratio [C/N0] based on the sensitivity of the receiver. The impact of the impedance bandwidth on the phase characteristics also need to be considered carefully while considering group delay effects, especially in wideband signal BOC waveforms such as the M code in modernized GPS and similar waveforms in Galileo. A good polarization axial ratio is an important factor in GNSS since it reduces multipath. This is because the signal generated by the first reflection off a multipath source is LHCP and this is reduced by an antenna with a good axial ratio.

The impedance bandwidth (BW) is determined by the maximum voltage standing-wave ratio (VSWR) defined by the symbol $S$.

$$BW = \frac{S - 1}{Q \cdot \sqrt{S}}$$  

(2.9)
In the above equation, \( Q_T \) is the total quality factor of the patch antenna [21] and \( S \) is defined in terms of the input reflection coefficient \( \Gamma \) as

\[
S = \frac{1 + |\Gamma|}{1 - |\Gamma|} \quad (2.10)
\]

The reflection \( \Gamma \) coefficient is a measure of the reflected signal at the antenna feed point before it is attached to the quadrature hybrid or branch-line coupler used for generating RHCP. It is defined in terms of the input impedance \( Z_{in} \) of the antenna and the characteristic impedance \( Z_0 \) of the transmission line feeding the antenna as given below:

\[
\Gamma = \frac{Z_{in} - Z_0}{Z_{in} + Z_0} \quad (2.11)
\]

The impedance bandwidth \( BW \) is normally specified as the frequency range over which \( S \) is less than 2. This corresponds to a power loss of 0.454 dB in the power delivered from the antenna to the receiver or to a return loss of –9.5 dB or to 11% of reflected power. The bandwidth of a microstrip antenna can be increased by reducing the Q factor as indicated in (2.6). The Q factor is proportional to the dielectric constant of the substrate and inversely proportional to its thickness. Hence the bandwidth can be increased by selecting a thicker substrate with a low-dielectric constant among the list of substrates listed in Table 2.1. However, this would increase the size of the microstrip patch antenna, which may not be desirable or allowable in certain commercial and military applications.

A good closed-form approximation for the impedance bandwidth of a patch antenna is [21]

\[
BW = \frac{16}{3\sqrt{2}} \frac{p}{\varepsilon_r} \left( \frac{1}{\varepsilon_r} \right) \left( \frac{b}{k_0} \right) \left( \frac{W}{L} \right) q \quad (2.12)
\]

where

\[
p = 1 - 0.16605 \left( k_0 W \right)^2 + \frac{0.02283}{560} \left( k_0 W \right)^4 - 0.0019142 \left( k_0 L \right)^2 \quad (2.13)
\]

\[
q = 1 - \left( \frac{1}{\varepsilon_r} \right) + \frac{2}{5\varepsilon_r^2} \quad (2.14)
\]

\( \varepsilon_r \) = radiation efficiency of the antenna. Expressions for \( \varepsilon_r \) are derived in detail in the next section.
Figure 2.11 shows the change in bandwidth as a function of the thickness of the dielectric substrate for three different values of the dielectric constant of the substrate: \(\varepsilon_r = 2.35\), \(\varepsilon_r = 6.0\), and \(\varepsilon_r = 9.2\). Notice that for a fixed thickness of the substrate, the bandwidth of the patch antenna decreases as the dielectric constant is increased; also for a fixed dielectric constant of the substrate the bandwidth increases as the substrate increases in thickness from 0.025\(\) to 0.3\(\).

While measuring the return loss of a GNSS antenna to determine its bandwidth using a network analyzer, it is important to note many commercial GNSS antennas already have a quadrature hybrid, such as a branch-line coupler, built directly across the antenna terminals and concealed inside the antenna package to obtain RHCP. Hence it is impossible to separate out the effects of this hybrid coupler on the antenna impedance. This prevents a measurement of the actual return loss of the antenna since the signal reflected from the input port of the antenna is diverted to the fourth port of the hybrid coupler, which has a 50-ohm matched termination. The hybrid coupler has a much broader bandwidth than the antenna; the signal reflected back from the antenna into the input terminal, where the return loss is being measured, would be negligible, resulting in a very flat response over a large frequency band. This would then lead to the erroneous conclusion that the antenna bandwidth is much broader than its true bandwidth. In cases where a hybrid coupler is built into the antenna module, a measurement of the antenna gain versus frequency is a more accurate measure of the true bandwidth of a GNSS antenna than the conventional definition using reflection loss at the input port.

In the case of an active GNSS antenna with a low-noise LNA built into the antenna package, the definition of antenna bandwidth becomes even more ambiguous since it is impossible to separate the frequency behavior of the antenna from the LNA. In such cases the only criteria for defining antenna performance becomes the gain/noise temperature or the G/T ratio as explained in Chapters 1 and 4.

![Figure 2.9](image.png)

**Figure 2.9** Variation in percentage bandwidth versus thickness and dielectric constant of the substrate. Resonance frequency = 1.5754 GHz.
2.2.4.4 Radiation Efficiency of an Antenna

The radiation efficiency \( e_r \) is defined by the ratio of radiated power \( P_r \) to the input power \( P_i \). The input power \( P_i \) is distributed between the radiated power \( P_r \), \( P_{sur} \) is the surface wave power propagated in the grounded dielectric substrate, \( P_c \) is the power dissipated in the metallization used for the patch antenna, and \( P_d \) is the power dissipated in the dielectric substrate. \( P_c \) is proportional to the square root of the conductivity of the patch metallization and \( P_d \) is proportional to the loss tangent of the dielectric substrate. For low-loss dielectric substrates with copper cladding, the dielectric loss \( P_d \) and conductor loss \( P_c \) are both very small so the radiation efficiency can be simplified and expressed as

\[
e_r = \frac{P_r}{P_i} = \frac{P_r}{(P_r + P_c + P_d + P_{sur})} \approx \frac{P_r}{(P_r + P_{sur})} \tag{2.15}
\]

Closed form expressions for \( P_r \) and \( P_{sur} \) has been derived [21, pp. 285].

\[
P_r = 40k_0^2 (k_0 h)^2 \left[ 1 - \frac{1}{\varepsilon_r} + \frac{2}{5\varepsilon_r^2} \right] \tag{2.16}
\]

\[
P_{sur} = 30\pi k_0^2 \varepsilon_r x_0^2 - 1 \left[ \frac{1}{\varepsilon_r} + \sqrt{\frac{1}{\varepsilon_r} - 1} \right] + k_0 h \left[ 1 + \frac{\varepsilon_r (x_0^2 - 1)}{\varepsilon_r - x_0^2} \right] \tag{2.17}
\]

where

\[
x_0 = \frac{\beta}{k_0} \approx 1 + \frac{1}{2} \left( \frac{\varepsilon_r - 1}{\varepsilon_r h} \right)^2
\]

\( x_0 \) is the normalized phase constant of the TM0 of the surface mode which has no cutoff frequency. The suppression of the surface wave mode will be discussed in greater detail in Section 2.2.8.1.

Figure 2.10 shows the variation in radiation efficiency of the square-patch antenna versus the dielectric constant and thickness of the dielectric substrate. Notice that the radiation efficiency is poor for thin substrates, but it improves rapidly as the substrate thickness increases; it then levels out for low-dielectric substrates such as for \( \varepsilon_r = 2.35 \) since these lower dielectric substrates do not generate much surface waves except for the fundamental TM0 mode. However, as the thickness increases for substrates with higher dielectric constants, as shown for \( \varepsilon_r = 6.0 \) and \( \varepsilon_r = 9.2 \) in Figure 2.10, the increase in surface wave radiation \( P_{sur} \) decreases the radiation efficiency as indicated in (2.15) when the thickness in increased beyond a certain level. This increase in surface wave radiation is not necessarily detrimental for GNSS.
applications since it boosts the gain at lower-elevation angles (makes the antenna beamwidth broader and lowers directivity), allowing the antenna to acquire low-elevation satellites and thereby improving PDOP.
2.2.4.5 Radiation Pattern and Axial Ratio of an RHCP Square Microstrip Antenna

The radiation pattern of an RHCP square-shaped microstrip antenna of length $L$ and width $W$ is equal to $a$ is placed on a very large (infinite) ground plane can be calculated by modeling the antenna as a combination of two orthogonal pairs of slot antennas with each pair of slots excited separately by a directly coupled probe, as shown in Figures 2.4(a) and (b). The first probe, Probe 1, excites the TM$_{10}$ mode in the two parallel slots of length $L = a$ parallel to the Y axis and are spaced at a distance $W = a$ along the X axis. The radiation pattern of the two parallel slots is linearly polarized with the electric field directed along the X axis parallel to the spacing distance $W$. The second probe, Probe 2, excites the TM$_{01}$ mode in the other pair of parallel slots of length $W = a$ oriented parallel to the X axis and spaced at a distance $L = a$ apart along the Y axis. The radiation pattern of this second parallel slot pair is also linearly polarized but with the electric field directed along the Y axis, orthogonal to the fields produced by the first pair of slots. RHCP is obtained by making the amplitude of the two orthogonal linearly polarized fields generated by the two slot pairs to be equal in amplitude but with a relative phase difference of 90°, with Probe 1 with 0° phase, and Probe 2 with −90° phase. If the voltage across either radiating slot is taken as $V_0$, the radiation fields for the pair of slots can be obtained by multiplying the radiation pattern of a single slot with an array factor to represent the parallel pair. For Probe 1 exciting the TM$_{10}$ mode, the component $E_\theta$ and $E_\phi$ of the far-field radiation is given by:

\[
(E_0)_{TM_{10}} = -j k_0 V_0 L \frac{e^{-j k_0}}{4\pi r} \cos \varphi F_1 F_2
\]  

(2.18)

\[
(E_0)_{TM_{01}} = -j k_0 V_0 L \frac{e^{-j k_0}}{4\pi r} \cos \theta \sin \varphi F_1 F_2
\]  

(2.19)

where

\[
F_1 = \sin c \left\{ \frac{k_0 b \sin \theta \cos \varphi}{2} \right\} \sin c \left\{ \frac{k_0 L \sin \theta \sin \varphi}{2} \right\}
\]  

(2.20)

and

\[
F_2 = 2 \cos \left\{ \frac{k_0 W \sin \theta \cos \varphi}{2} \right\}
\]  

(2.21)

Similarly, for Probe 2 exciting the TM$_{01}$ mode, the component $E_\theta$ and $E_\phi$ of the far-field radiation is given by
\[
(F_0)_{TM_{01}} = jk_0 V_0 W \frac{e^{-ikz}}{4\pi r} \sin \varphi F_1^l F_2^l
\]

\[
(F_\varphi)_{TM_{01}} = jk_0 V_0 W \frac{e^{-ikz}}{4\pi r} \cos \theta \cos \varphi F_1^l F_2^l
\]

where

\[
F_1^l = \sin e \left\{ k_0 \frac{b \sin \theta \sin \varphi}{2} \right\} \sin e \left\{ k_0 W \sin \theta \cos \varphi \right\}
\]

and

\[
F_2^l = 2 \cos \left\{ k_0 L \sin \theta \sin \varphi \right\}
\]

Since we are considering a square-shaped patch antenna, \(W = L = a\) in the above equations. Signals for the TM_{10} and TM_{01} modes generated in the patch antenna are parallel to the \(\hat{x}\) and \(\hat{y}\) axes, respectively. RHCP is generated when these two orthogonal signals are combined with equal amplitudes but in phase quadrature such that their ratio \(\frac{E_{RHCP}}{E_{Y}} = j\); this is achieved through a branch-line coupler.

By examining the equations above the following features are noticed:

1. At zenith \(\theta = 0\) and, \(\cos \theta = 0\), \(\hat{E}_{RHCP} = [C]\left\{ \hat{\theta} \sin \varphi + j \cos \varphi + \hat{\varphi} \cos \varphi - j \sin \varphi \right\}\)

\(\hat{E}_{RHCP} = [C]\hat{e}_x\) is the RHCP electric far field from the patch antenna expressed in spherical coordinates with the wave traveling in the +r direction and.

\(C = k_0 V_0 a \frac{e^{-ikz}}{4\pi r} \hat{e}_x = \hat{\theta} \left[ \sin \varphi + j \cos \varphi \right] + \hat{\varphi} \left[ \cos \varphi - j \sin \varphi \right]\)

is defined as the RHCP complex unit vector [29, p. 59]; hence at zenith the patch antenna is purely RHCP and the CP axial ratio is 1. Similarly, for a wave traveling in the +r direction we can also define (using this same type of derivation) an LHCP complex unit vector \(\hat{e}_y = \hat{\theta} \left[ \sin \varphi + j \cos \varphi \right] + \hat{\varphi} \left[ \cos \varphi - j \sin \varphi \right]\). The LHCP electric far field from the patch antenna expressed in spherical coordinates can now be represented as \(\hat{E}_{LHCP} = [C]\hat{e}_y\). Two parameters are often used to define purity of CP. The most commonly used parameter is the axial ratio “R” for CP defined as \(R = \frac{\hat{E}_{RHCP}}{\hat{E}_{LHCP}} + \frac{\hat{E}_{LHCP}}{\hat{E}_{RHCP}}\); a second parameter used less often is the polarization ratio for CP \(\rho_c = \frac{\hat{E}_{RHCP}}{\hat{E}_{LHCP}}\). Figure 2.13 shows the measured RHCP (i.e., principal polarization) and LHCP (i.e., cross-polarization) radiation patterns of a microstrip patch antenna measured in the upper hemisphere down to
an elevation of $-30^\circ$ below the horizon at a frequency of 1.5754 GHz. A 51" diameter rolled edge ground plane was used in this pattern measurement; this ground plane is shown in Figure 2.12(d). The edges of the ground plane are rolled inwards below the top surface of the ground plane to prevent the signals diffracted from the edges from affecting the radiation pattern of the antenna in the upper hemisphere.

2. We notice that for a microstrip patch antenna located on an infinitely large ground plane, the horizontally polarized $\hat{\phi}$ component $(E_\phi)_{TM_{10}}$ and $(E_\phi)_{TM_{01}}$ of the far-field radiation for the TM$_{10}$ and the TM$_{01}$ modes are both zero at the horizon since $\cos \theta = 0$. Hence the far-field radiation from the patch antenna at the horizon is linearly polarized with just the vertically polarized $E_\theta$ component. This can also be concluded by noticing that the horizontal $E_\phi$ component being tangential to the large conducting ground plane needs to be zero. The axial ratio for CP of the antenna is $\infty$ at the horizon. Hence the purity of CP and axial ratio of a patch antenna on an electrically large ground plane is elevation angle-dependent with good RHCP obtained only near zenith but with progressive degradation in axial ratio as the horizon is approached. The polarization becomes vertically (i.e., linearly) polarized at the horizon with no perceptible difference between RHCP and LHCP.

2.2.4.6 Half-Power Beamwidth Gain and Directivity of an RHCP Microstrip Antenna

The half-power beamwidth (HPBW) of a receiving antenna is defined as the angular width between directions where the power of the received radiation is decreased by 3 dB. A more appropriate performance measure for a GNSS antenna is its minimum

Figure 2.12 Measured radiation pattern of a GNSS patch antenna on two different types of finite size ground planes. (a) Measured pattern of 4-foot square-shaped ground plane, (b) measured pattern on rounded edge ground plane, (c) diagram of 4-foot square ground plane, and (d) actual rounded edge ground plane. Resonance frequency = 1.5754 GHz.
gain beam width or the angular region in the upper hemisphere above the minimum masking elevation angle where the antenna is able to meet the minimum gain needed to acquire GNSS satellites when used with a specific receiver (see Figure 2.2). The minimum gain requirement that the antenna is required to meet to acquire the satellite was discussed in Chapter 1. The HPBW $\theta_E$ in the E plane where $\phi = 0$ and the HPBW $\theta_H$ in the H plane where $\phi = 90^\circ$ of a single-layer patch antenna can be determined from its dimensions [21, p. 276–277].

$$\theta_H = 2\arcsin\left\{\frac{1}{2+k_0L}\right\}^{\frac{1}{2}} \quad (2.26)$$

$$\theta_E = 2\arcsin\left\{\frac{7.03}{3k_0^2W^2+k_0^2h^2}\right\}^{\frac{1}{2}} \quad (2.27)$$

Figure 2.13  GNSS dual-band RHCP microstrip antennas with direct contact feed probe. (a) Feed connected to top patch, and (b) feed connected to bottom patch; inverted configuration.

Since we are considering a square-shaped patch antenna $W = L = a$, where $a$ is the dimension of the patch and $b$ is the thickness of the dielectric substrate. The beamwidth of the antenna can be increased by decreasing the size of the patch antenna by reducing $a$; this can be accomplished by selecting a substrate with a high dielectric constant. This helps the antenna to meet the minimum gain beam width required to acquire GNSS satellites at even at lower-elevation angles. A reduced size GNSS antenna is also a requirement for many commercial and military applications, especially for avionics. The smaller size, however, needs to be balanced against the resulting reduction in bandwidth and decrease in antenna gain
and directivity. The resulting reduction in bandwidth can be recovered by either increasing the thickness of the dielectric substrate or by using stacking patch antennas tuned to different frequencies on top of each other as discussed later in this chapter.

The directivity of the antenna is a measure of its directional properties compared to that of an isotropic antenna and is defined as the ratio of the maximum power density in the main beam direction to the average radiated power density. The directivity $D$ of the patch is expressed as

$$D = \frac{\left( \frac{r^2}{2\eta_0} \left( |E_\theta|^2 + |E_\phi|^2 \right) \right)}{P_r \frac{\eta_0}{4\pi}}$$

(2.28)

where $P_r$ is the radiated power, $\eta_0 = 120 \, \pi$, and the radiated fields $E_\theta$ and $E_\phi$ were defined earlier in (2.22) through (2.26). A simple approximate expression for the directivity $D$ of the patch antenna is given by

$$D = \frac{4(k_o a)^2}{\pi \eta_0 G_r}$$

(2.29)

where $G_r$ is the radiation conductance of the antenna and $G_r = \frac{1}{R_r}$, where $R_r$ is the radiation resistance defined earlier. The gain $G$ of the antenna is defined as $G = e_r D$ where $e_r$ is the radiation efficiency of the antenna. Gain is always less than the directivity because the efficiency $e_r$ is in the range $0 < e_r < 1$.

2.2.4.7 Finite Ground Plane Effects on the Antenna Pattern of Microstrip GNSS Antennas

The interaction between a microstrip antenna and a finite-size ground plane on which it is mounted is a complex problem with serious consequences on GNSS performance. This topic is discussed in greater detail in Chapter 5, including methods used for both the analysis and the mitigation of such effects. The equations given above for the radiation patterns should be considered as valid only for an infinitely large ground plane and dielectric substrate. A finite ground plane has significant effects on the radiation pattern of a GNSS antenna from diffraction caused from the edges of the ground plane. These diffracted signals cause ripples to occur in the antenna pattern at higher-elevation angles close to zenith, affects the RHCP gain at lower elevations by changing the axial ratio, and also creates backlobes that can degrade the front-to-back ratio of the pattern and make the antenna susceptible to multipath and interference. Approximate expressions to account for these ground plane effects have been provided [21, pp. 293–296].

Figure 2.12 shows the radiation pattern of a microstrip patch antenna measured on two different types of ground planes, which illustrates how the size and shape of the ground plane influences the antenna pattern. Figure 2.12(a) shows
measured radiation pattern of the microstrip antenna at a frequency of 1.5754 GHz when it is placed at the center of a 4-foot square planar ground plane. The cross section of this 4-foot square ground plane is shown in Figure 2.12(c). The large ripples seen in the main beam pattern of the antenna shown in Figure 2.12(a) are caused by straight edge diffraction from the ends of this planar ground plane. Figure 2.12(b) shows measurements made on the same antenna when it is placed at the center of a 51" diameter rolled-edge ground plane whose edges have been rolled underneath the ground plane to reduce the impact of diffraction effects on the antenna pattern in the upper hemisphere. A picture of the rolled-edge ground plane is shown in Figure 2.12(d). The main beam is broader and smoother since the diffraction effects have been reduced due to rolling the edges. Notice the neither of these two ground planes is able to suppress the antenna backlobes; this would require the use of more sophisticated ground planes designs such as choke ring ground planes, electronic bandgap ground planes, and resistivity tapered ground planes designed specifically to mitigate the diffraction effects from the edges of the ground plane. The design of these special ground planes are discussed in greater detail in Chapter 5.

2.2.5 GNSS Dual-Band Stacked Microstrip Patch Antennas

Antennas needed for modernized GPS, Galileo, and GLONASS may need to operate in two or even three separate frequency bands for providing the highest measurement accuracy. Modernized GPS applications, for example, will require the antenna to have a minimum operating bandwidth of 20 MHz for civilian use and 24 MHz for military applications (for GPS M code) in each frequency band; the antenna also needs to be RHCP to allow optimum reception of the satellite signals. The instantaneous bandwidth needed to cover the entire frequency range of modernized GPS, Galileo, and GLONASS extending from 1166 MHz (the lower end of the L5 band) to 1607 MHz (the upper end of the GLONASS G1 band) would be 31.8\%, covering a frequency span of 441 MHz. It would be difficult to meet this wide instantaneous bandwidth with a single microstrip patch antenna, which being a resonant device with a high Q generally has a bandwidth of just a few percent, as shown in Figure 2.9. Since these antennas also need to be small in size for most GNSS applications, substrates with high dielectric constants are frequently used in their construction, which further restricts their bandwidth. However each specific GNSS band that the patch antenna is needed to cover is no more than 24-to 30 MHz wide; one method of circumventing the large instantaneous bandwidth problem is by integrating two or more narrowband microstrip antennas with each antenna covering only the required bandwidth of 4 MHz with a performance similar to that of a notched passband filter. This technique also provides the added advantage of avoiding out-of-band interference with each antenna acting as its own filter. However an integrated feeding technique is needed to connect the two or more patch antennas operating in different frequency bands to a common GNSS receiver that is often used to process signals in the various receiving bands. One popular technique that has been used to obtain dual-band performance is by “stacking” multilayer resonant patch antennas vertically, one on top of the other with the upper patch resonating at the higher frequency and the lower patch at a lower-frequency, as shown in Figure 2.13. The stacked patch antennas are all fed with a common feed and polarization network that is connected to the GNSS receiver. These stacked
patch designs have been described in several textbooks [33, 22]. Stacking also prevents increasing the lateral dimensions of the patch antenna, a key requirement in many GNSS applications such as avionics and military systems.

Two feed techniques have been used to feed multilayer patch antennas with the same set of two orthogonally placed probes. The first method is called the top-feed technique and is shown in Figure 2.13(a). The inner conductor of each feed probe passes through the bottom patch without making electrical contact; a small circle of the lower patch is removed for this purpose. The inner conductor proceeds through the dielectric substrate of the top patch and is then soldered to the top patch. The location of the two probes within the patch is optimized to provide a fairly good impedance match at both frequency bands. The larger bottom patch acts as a ground plane for the smaller top patch; the upper patch when resonant at the higher frequency has negligible reactance effect on the bottom patch effect and vice versa. Since there is strong electromagnetic coupling between the top and bottom patches, their resonant sizes can be determined accurately only through the use of advanced electromagnetic codes listed in Table 2.2. The HFSS code (described in Section 2.3.5.1) has been used for designing a dual-band patch antenna.

A second direct feed method that is less commonly used is the bottom-feed technique shown in Figure 2.13(b). Here the feed probe is connected only to the lower-band patch antenna at the bottom of the stack and the higher-frequency patch at the top is parasitically coupled to the bottom fed patch. The bandwidth of these bottom probe-feed stacked patch antennas can also be increased further by using the hi-lo stacked patch design obtained by selecting the appropriate substrate materials for the top and bottom patch antennas [28, pp. 56–67]. In this design the lower patch antenna element is mounted on a high dielectric constant substrate and the upper patch is mounted on a substrate of foam with \( \varepsilon_r \sim 1.07 \). This design would cause the size of the patch antenna to increase. However, this technique does not appear to have been tried for building GNSS antennas. Even larger bandwidths can be obtained by using aperture coupled, dual-band stacked patches [30]. Triple-band antennas have been developed by stacking up to three patch antennas and by using aperture-coupling techniques; these will be discussed under multiband GNSS antennas in Chapter 3.

2.2.5.1 Design of a Dual-Band Microstrip Antenna through HFSS Computer Simulations

As mentioned earlier there are no convenient CAD tools currently available to simplify the design of a dual-band stacked patch antenna with common feed probes for both bands; this is due to the complex coupling between the top and bottom patches. This can only be done accurately by using one the advanced computer design codes listed in Table 2.2.

In this section we will describe the design of a dual-band stacked patch antenna obtained from simulations using the HFSS computer code. A picture of the antenna that was designed and built is shown at the left in Figure 2.14(a); the corresponding HFSS simulation model of this antenna is shown at the right in Figure 2.14(b). The dielectric substrate used for this antenna was Rogers TMM4 with a dielectric constant of 4.5 and loss tangent of 0.002. The feed design used was the top-feeding technique shown earlier if Figure 2.13(a). Two direct-contact top-feed probes are
used to feed both the L1 patch antenna at the top and the also the L2 patch antenna at the bottom of this stacked structure. The top L1 patch is 1.68" square and the bottom L2 patch is 1.94" square. The top and bottom substrate layers are 0.150" and 0.3" in thickness, respectively. The size of the truncated dielectric substrate is 2.75". The return loss at the two frequency bands, the Smith chart showing dual resonances as calculated by the HFSS program, are shown in Figures 2.15(a) and (b), respectively. The measured return loss for this antenna is shown in Figure 2.16 and agrees well with the HFSS simulations shown in Figure 2.15(a); the measured RHCP and LHCP antenna patterns and gain at 1.5754 and 1.2276 GHz, and the center band frequencies of the GPS L1 and L2 bands are shown in Figure 2.17.
2.2.5.2 Coplanar Dual-Band GPS Patch Antennas

A second method for designing a compact, dual-band RHCP microstrip antenna for the L₅/L₁ bands of GPS is shown in Figure 2.18. This design uses a thin circular ring antenna resonant in the L₅ band (the lower-frequency band) that is parasitically coupled to a concentrically located circular patch antenna resonant in the L₁ band.
higher-frequency band $L_1$ [31]. Schematic sketches of the top and side view cross section of this dual-band antenna is shown in Figure 2.18. Pictures of the top and side views of this antenna are shown in Figure 2.18(b). This compact antenna meets the ARINC 743 height and width requirement needed for avionics; the ARINC cross-sectional requirement was shown earlier in Figure 2.3. The centrally located circular patch in this design is fed by two or more direct-contact feed probes for generating RHCP in both patch antennas. In this design antennas operating in both frequency bands are coplanar, which allows them to share a common dielectric substrate including a compound substrate consisting of two or more different substrate materials with different thicknesses and dielectric constants.

The new design will allow increased flexibility in controlling the size of the antenna elements and provide wide antenna patterns with good gain coverage at the lower-elevation angles in both the frequency bands. Good gain at lower elevations is particularly important in the $L_5$ band to better withstand potential interference from other in-band ARNS transmissions. Both antennas are RHCP although only the central element has feed probes connected to it with the outer element being parasitic. The antenna dimension without the radome is 1.56" square with a height of 0.6". It uses a compound substrate: the top substrate layer is 0.2" thick and has a dielectric constant of 12.78 (Rogers TMM 13i) and the bottom substrate is 0.4" thick and is made from a ceramic-type material with a dielectric constant of 30 manufactured by the Emerson & Cummings Corporation. Figure 2.19 shows the measured return loss of this antenna. Figure 2.20 shows the measured RHCP and LHCP antenna patterns at 1.5754 GHz and 1.176 GHz—the center-band frequencies in the GPS $L_1$ and $L_5$ bands.

Another coplanar design for a circularly polarized single-layer dual-band microstrip antenna has also been proposed by Fan and Rahmat-Samii [32]. This uses a square patch with stubs attached on opposite ends as well as four slots parallel to the edges of the main square-patch antenna with switches installed in each slot.
92 FRPAs and High-Gain Directional Antennas

The design is complex, and also the axial ratio appears to be poor. This design does not allow the antenna to operate simultaneously in both bands since the antenna needs to be switched from one band to the other through the PIN diodes.

2.2.6 Feed Techniques for Generating RHCP in GNSS Microstrip Antennas

The polarizing feed network used for generating RHCP in a GNSS patch antenna is another important feature in its design. It has an impact on many GNSS
performance parameters such as CP axial ratio bandwidth, the PCV relative to look angle and frequency, and RHCP antenna gain and pattern symmetry. Different feeding techniques have been developed varying from a simple single feed for low-cost, narrowband applications to broadband multilayer aperture coupled networks covering all three frequency bands of modernized GPS. They can be grouped into four broad categories: direct-contact probes, aperture-coupling, edge-coupling, and proximity-coupling techniques; these are illustrated in Figures 2.21, 2.22, and 2.23(a) and (b), respectively. Each feed technique offers some advantage in terms of performance, complexity, and cost that is best suited for a specific application. The probe and aperture coupling techniques are two that are most popular. In the

Figure 2.21 Four direct-contact feed probes for generating low-axial ratio RHCP in GNSS microstrip antennas.

Figure 2.22 Aperture-coupled dual-band microstrip antennas. (From [33] ©2008 IEEE.)
direct-probe feeding method the RF signals received by the patch antenna from the GNSS satellites are fed to the receiver using one or more coaxial probes in direct contact with one or more patch antennas. Direct-contact probe-feed techniques are popular because of their simplicity. The use of one, two, or four probes are most common in many GNSS applications. They can generate RHCP in square, circular, and annular ring patches. Figure 2.5(a through c) shows pictures of RHCP GPS microstrip antennas for these three different geometries fed either by two or four direct-contact probes.

Two- and four-feed probes, shown in Figures 2.24 and 2.21, respectively, are more popular for general use since they are less complex and provide adequate performance for many applications. The two-probe technique was discussed in detail earlier in this chapter. The PCV variation can be reduced further by using four-feed probes [34]. This feed design is used in high-quality GNSS antennas and provides

Figure 2.23 (a) Edge- and (b) proximity-coupling techniques for GNSS microstrip antennas.

Figure 2.24 Microstrip branch-line hybrid coupler for generating RHCP in a two-probe feed microstrip antenna.
the best compromise between simplicity and good pattern symmetry and low PCV; it will be discussed in detail later in this chapter.

Up to eight probes (called an N-point feed) has been used by Trimble recently to develop high-quality geodetic antennas with very good phase center stability such as in their Zephyr antenna [35, 36]. Using more probes helps to suppress higher-order modes and reduces cross-polarization levels by preserving modal purity. This is achieved by better adherence to the optimum phase relationship between the probes required for exciting the fundamental mode. It produces greater symmetry in the azimuth pattern with a more stable phase center and lower CP axial ratio. The disadvantage of a multipoint feed is the reduction in gain due to losses in the feed network and increased cost and complexity.

In the aperture coupling technique shown in Figure 2.22, the microstrip feed line that is connected to the GNSS receiver is on a separate dielectric substrate from the patch antenna; it is in close physical proximity but not in direct physical contact with the patch. The transfer of RF power between the patch antenna and the microstrip line is through electromagnetic coupling from intervening resonant slots cut into the ground plane. The advantage provided by these noncontact techniques is the vast improvement in bandwidth over what direct-contact probe feeds provide. Crossed slots for aperture coupling can produce very wide bandwidths that can cover multiple GNSS bands by using stacked patch antennas excited by crossed resonant slots. Dual-band [37] and triple-band [33] GNSS antennas using these aperture feeding techniques have been built and will be discussed in Chapter 3 under multiband antennas.

Edge coupling is another type of a direct-contact feed where two microstrip transmission lines are connected to the patch antenna at its edges, as shown in Figure 2.23(a) [23, pp. 85]. The other ends of the microstrip lines are connected to a coplanar broadband hybrid on the same substrate to obtain a 90° phase difference between the two ports. The advantage is that the feed network can be colocated with the patch antenna on the same printed circuit board, obviating the need for soldering as in the direct feed probes and reducing the manufacturing cost. The disadvantage is the large lateral size needed to support the feed around the radiating patch. Proximity coupling, shown in Figure 2.23(b), is similar to aperture coupling; the microstrip feed line that is connected to the GNSS receiver is on a separate dielectric substrate from the patch antenna but is in close physical proximity to but not in direct contact with the patch.

We will now consider the design of some of these feed techniques in greater detail.

### 2.2.6.1 RHCP Feed Design Using Two Direct-Contact Probes

The design of a square-shaped microstrip patch antenna with two direct-contact feed probes was covered in a previous section. This type of feed technique provides several advantages:

1. The quadrature hybrid to which these probes are connected has a very wide bandwidth, sometimes covering a full octave. Hence RHCP with a low axial ratio can be obtained over a wide range of frequencies with a limitation on bandwidth imposed only by the impedance variation with frequency of
the feed probe, which is at a fixed location inside the patch and therefore provides an optimum match only at one frequency.

2. The feed network is isolated from the patch antenna by the ground plane, resulting in a good front-to-back ratio especially when the ground plane is designed to reduce edge diffraction effects as explained in Chapter 5. This decreases the susceptibility of the GNSS antenna to multipath.

3. The coaxial probes are normal to the surface of the patch and do not increase the lateral dimensions of the antenna; this is a huge advantage when multiple patch antennas need to be located close to each another within the small aperture of an adaptive array needed for antijam applications. The available space for adaptive antenna arrays in tactical military aircraft is often smaller than a wavelength in diameter at GPS frequencies for miniaturized arrays. These space restrictions only allow antenna elements with only very small lateral dimensions including the feed networks to be used in such small antenna arrays.

4. The quadrature hybrid, a schematic sketch of which is shown in Figure 2.24, can be used for extracting both RHCP and LHCP signals from the patch antenna using Ports 1 and 4, respectively.

5. Direct-contact probe feed techniques also avoid misalignment of multiple dielectric layers encountered in aperture coupled patches.

A key component in the feed network for generating RHCP is the quadrature hybrid, a schematic cross section of which is shown in Figure 2.24 and whose scattering matrix is given by:

\[
[S] = \frac{1}{\sqrt{2}} \begin{bmatrix}
0 & j & 1 & 0 \\
 j & 0 & 0 & 1 \\
1 & 0 & 0 & j \\
0 & 1 & j & 0 \\
\end{bmatrix}
\]

(2.30)

For generating RHCP, Port 1 is connected to the GNSS receiver and Port 4 is terminated in a matched load. Alternatively, a cross-polarized LHCP output signal can be obtained from Port 4 when it is not terminated in a matched load; this is useful when obtaining both RHCP and LHCP signals from the antenna element for use in dual polarized adaptive antenna arrays to be discussed in Chapter 4. Ports 2 and 3 are connected to the two orthogonally placed probes in the patch antenna as shown in Figure 2.24. The signal is evenly divided in amplitude but in phase quadrature at Ports 2 and 3, with a phase of 0° from Port 2 and a phase of −90° from Port 3. For obtaining a 3-dB split between these two ports with the reference impedance \( Z_0 \) generally set to 50 ohms, the shunt branches of the hybrid \( Z_S = Z_0 \) and the through branches \( Z_T = \frac{Z_0}{\sqrt{2}} = 35.4 \) ohms. The lengths of the branches are all a quarter wavelength at the center design frequency. Port 4, which is the isolated port if not used for dual polarization application, is generally terminated in a matched load to absorb any imbalance between Ports 2 and 3. Any signals reflected as a result of an impedance mismatch between the patch antenna and Ports 2 or 3...
will be returned to Port 4 and absorbed by the match termination. When a quadrature hybrid is connected to the patch antenna and integrated within the antenna package, a feature most common in commercial GNSS antennas, the reflected signals from the antenna ports are not reflected back into the input Port 1. Therefore a relative measurement of the forward and reflected powers is not possible, and the bandwidth of these antennas cannot be estimated by measuring its VSWR or return loss. Some of these antennas can have very poor gain and bandwidth but still show deceptively low return loss as a function of frequency indicating good bandwidth.

2.2.6.2 RHCP Feed Design Using Four Direct-Contact Probes

One method of increasing the bandwidth of a patch antenna is by increasing its thickness as described in the previous sections. However, due to the unbalanced and asymmetrical nature of the dual probe feed arrangement for generating RHCP, this can result in the generation of unwanted higher-order modes inside the patch antenna. The mode closest to the desired \( \text{TM}_{01} \) and \( \text{TM}_{10} \) modes is the \( \text{TM}_{21} \) mode, which causes coupling between the dual feed probes in the patch antenna and affects the amplitude and phase balance between the probes even when the antenna is connected to the quadrature hybrid. This imbalance increases the cross-polarized LHCP signal and distorts the radiation pattern. These effects can be reduced by replacing the dual feeds by a balanced four-probe feed system as shown in Figure 2.21. In this feed arrangement, two additional probes have been added to the two original ones and feeding all four terminals with equal amplitude but 90° out of phase from each other. A new feed circuit consisting of two 90° hybrids that are combined with a 180° hybrid. The four probes have equal amplitudes but phases

---

Figure 2.25 Toko/America’s GPS ceramic microstrip antenna with a single-feed probe. GPS L1 band; C/A code. (From: [39]. Printed with permission from Toko/America.)
of 0°, −90°, −180°, and −270° varying in a counterclockwise direction to generate RHCP. These two orthogonally placed pairs of probes have opposite phases but equal amplitudes and result in a cancellation of the unwanted TM21 modal fields. Four coaxial probes provide good symmetry in the azimuth plane of both the phase and the radiation pattern of the antenna. This type of feed also provides the low phase and group delays [34]. The disadvantage of the four-probe feed method besides the increased complexity is a lower antenna gain than obtained with the dual probe feed network because of the losses contributed by the additional hybrid components needed in the feed network.

Difficulties can arise in impedance matching when a thick dielectric substrate is used for increasing the bandwidth and gain. The extension of the center conductor of a directly coupled feed probe through the dielectric substrate introduces a series of inductive reactance, which is proportional to the substrate thickness. This increased inductance can limit the impedance bandwidth of the patch antenna. Modified types of coaxial probes for feeding patch antennas have been designed recently to compensate for the increase in inductance by using capacitive coupling between the probe and the patch antenna, thereby avoiding direct contact between the two. These techniques are more similar to proximity coupling techniques although the probes are embedded in the substrate beneath the patch antenna. Various methods have also been developed to tune out the probe inductance to increase the bandwidth [30, 31]. They include etching a small hole in the patch at the top around the probe, or attaching a small conducting strip to the top of the probe and then coupling it to the patch either from above or from below, or by using a proximity-coupled L-shaped probe. In an L-shaped probe the horizontal part of the probe runs underneath the patch and provides capacitive coupling to it [40]. Two orthogonal L-shaped probes have been used recently in a broadband patch antenna covering all three of modernized GPS frequency bands [41].

2.2.6.3 Single Direct-Contact Feed Probe Design for RHCP GNSS Patch Antennas

A single direct feed probe placed along the diagonal axis of the patch antenna at a suitable distance from the center is often used in low-cost and compact antennas produced for the mass market that operate only over a single GNSS band. It is especially popular with high-permittivity ceramic substrates [38, 39]. Figure 2.27 shows a popular, single-probe fed GPS antenna (part number DAKC1575MS74T) manufactured by Toko America Inc. Use of a single direct-contact probe for generating CP removes the need for an external phasing network and greatly reduces cost, complexity, and size. However, there are two penalties paid for this reduction in complexity:

1. It generates the largest variation in the phase center with elevation and azimuth angle, resulting in an increase in the phase center error ellipsoid [34, 38].
2. It has a very narrow axial ratio CP bandwidth. The PCV variation results from the generation of higher-order modes in the patch antenna due to the asymmetric excitation; this increases the cross-polarized levels contributing to the pattern asymmetry [34, 38].
The narrow RHCP axial ratio bandwidth makes this design suitable for use only over a single GNSS frequency band, generally over only the C/A code in the GPS L₁ band. The single-probe patch antenna is able to generate RHCP by exciting mutually orthogonal TM₁₀ and the TM₀₁ modes with the same amplitude but with a phase difference of 90°. This is done by several different methods illustrated in Figures 2.26 (a through d). A more detailed description of these various single-feed probe designs for generating RHCP can be found in two recent books [30, pp. 319–339] and [23, pp. 37–49]. Two single-probe designs for a GPS patch are discussed below; the resonance frequency of the antenna is 1.5754 GHz.

One of the techniques used is by altering the shape of the perfectly square patch to “nearly square” by changing the aspect ratio of width/length as shown in Figure 2.26(a). When the patch antenna is perfectly square, it will excite the orthogonal TM₀₁ and the TM₁₀ modes with identical amplitudes and in phase; the phase centers of these modes also coincide and the far field will have a linear polarization directed diagonally. However, a slight departure from a perfectly square shape causes the two spatially orthogonal modes, the TM₁₀ and the TM₀₁ modes,

![Figure 2.26](image_url)

**Figure 2.26** Different types of single-feed probe techniques for generation of RHCP in GNSS microstrip antennas: (a) diagonally fed nearly square patch, (b) square patch with diagonal corners cut off, (c) square patch with indentations with a diagonal feed, and (d) square patch with a diagonal slot.
to have different resonant frequencies that are close to each other. The resonance frequency of the $TM_{01}$ mode $f_2$ is higher than the desired signal frequency whereas the resonant frequency of the $TM_{10}$ mode $f_1$ is lower than the desired signal frequency $f_0$. The feed location of the probe is next located at a point where the antenna is excited at the desired frequency $f_0$ in between the resonance frequencies $f_1$ and $f_2$ so that the two modes have equal amplitudes but a phase difference of $+45^\circ$ or $-45^\circ$ with respect to the feed point to meet the conditions needed for obtaining RHCP. However, these conditions of amplitude and phasing can only be met exactly over a narrowband where the frequency ratio of the two orthogonal modes is $1.01 \left( \frac{f_2}{f_1} \right) 1.10$. The axial ratio bandwidth is small generally around 1%. To illustrate this method for generating RHCP, consider a nearly square microstrip patch on a 0.254-cms thick dielectric substrate with a dielectric constant $\varepsilon_r = 6.0$ and loss tangent $\delta = 0.0018$. If the patch was square, the dimensions of each side would be equal to 3.8245 cms to obtain resonance at a frequency of 1.5754 GHz, and RHCP could be generated by using two feed probes located at orthogonal positions in the patch as illustrated in the previous section. To obtain RHCP with a single probe at the same design frequency of 1.5754 GHz, the shape of the patch antenna can be made nearly square with a width along the $x$ axis equal to $L + c = 3.8757$ cms and length along the $Y$ axis equal to $L = 3.8245$ cms or an aspect ratio of 1.0138 corresponding to a width extension $c = 0.0512$ cms. The feed probe is located along the diagonal at $X_p = 0.3039$ cms and $Y_p = -0.2783$ cms. The best RHCP is obtained at a frequency of 1.5634 GHZ, slightly below the desired frequency.

A second popular technique for generating RHCP with a single feed, shown in Figure 2.26(b), is to change the shape of the patch by truncating the two diagonal opposite corners of the square-patch antenna by a length $c$. This again creates a pair of diagonal modes that can be adjusted to have identical magnitudes and a 90° phase difference between the modes. By placing the feed probe at an appropriate distance from the center along the central $x$ axis it is possible to excite each of these diagonal modes to have equal amplitude so as to provide the best RHCP as well as a good impedance match. To illustrate this method consider the same dielectric substrate used in the previous example of the nearly square patch. The dimension of the patch is 3.8245.cms square except that two opposite corners are truncated by distance $c = 0.3129$ cms. The feed is located at $X_p = 0.3650$ cms and $Y_p = 0.0$. The best RHCP is obtained at 1.5634 GHz, again slightly below the desired frequency of 1.5754 GHz. Two other methods for obtaining RHCP with a single feed probe are shown in Figures 2.26(c) and (d). These methods change the geometry of the square patch by inserting indentations or notches or by cutting a diagonal slot across the patch.

2.2.6.4 Aperture Coupled Feed Design for RHCP GNSS Patch Antenna

In this feed technique, shown in Figure 2.22, the radiating patch and the microstrip feed line are separated by a common ground plane that they share. Coupling between the patch and the feed line is made through a crossed slot aperture in the ground plane that is located directly below the center of the patch. This type of feed technique is sometimes also called aperture stacked patches (ASPs); for GNSS
applications dual-polarized ASPs are of interest [33, 37]. This feed configuration has the ability to isolate the patch radiator from the feed line by virtue of the common ground plane that they share. This minimizes the effects of the spurious radiation from feed on the antenna pattern of the patch and also allows the design of the feed and the antenna to be optimized relatively independently of each other. The resonant frequency is primarily determined by the size of the square-patch antenna but the resonant slot length of the aperture exciting the patch may also have a secondary effect on the resonance frequency. The resonant frequency of the patch may show a slight decrease as the length of the aperture increases. A decrease in the length of the aperture may also decrease the level of coupling between the microstrip feed line and the patch antenna. The dielectric constant of the antenna substrates are also a factor in determining antenna performance. Increasing the permittivity of the feed substrate has minimal impact on the resonant frequency, but the coupling between the feed line and the patch antenna is increased as a result of the aperture now being electrically larger than its physical dimensions. If the thickness of the feed substrate is increased the level of coupling to the patch antenna is decreased, but the resonant frequency is not affected.

An aperture feeding technique used for generating RHCP is a crossed-slot aperture fed by microstrip lines from a parallel feed configuration consisting of three Wilkinson power dividers shown in Figure 2.22. Two orthogonal linear modes are excited by the crossed slot aperture that are of equal amplitude but with a 90° phase shift for achieving CP in the patch antenna. The 90° phase shift between the two slots is produced by extending one arm of the Wilkinson power divider by a quarter wavelength. The two Wilkinson power dividers provide diametrically opposite ends of the crossed slots with a 180° phase difference by being connected to a third Wilkinson power divider. This combination provides equal amplitudes but with a phase progression of 0, 90°, 180°, and 270° across the four ends of the pair of slots to generate RHCP. This type of feeding technique has a very broad impedance and axial ratio bandwidths and allows the stacking of two square-patch antennas above the crossed-slot aperture that are resonant at different frequencies to obtain multiband performance.

Aperture-coupled feeding techniques also present two major disadvantages for GNSS applications. The first is the backward radiation into the lower hemisphere by the coupling aperture cut into the ground plane. The backward radiation degrades the front-to-back ratio (FBR) of GNSS antennas; this makes it more susceptible to multipath effects. This is apparent from the measured patterns of the two GNSS antennas that have been built using this type of feed [33, 37]. The measured backlobes in a GPS dual-band aperture coupled patch antenna is shown in Figure 2.27 [37].

Techniques for reducing backward radiation in both polarizations have been proposed by Waterhouse [18, pp. 139–145]. One of the methods suggested is to use a printed cross-back patch below the resonant crossed slots used for exciting the patch antennas. This microstrip patch element placed behind the microstrip-fed aperture acts as a reflector to substantially reduce backward radiation. By using this method Waterhouse has been able to reduce FBR typically from between 10 to 14 dB to more than 20 dB. However, additions to the feed circuit greatly add to the size, cost, and complexity. The second disadvantage of aperture coupling is the increase in lateral dimensions of the antenna resulting from the feed structure
especially for the parallel feed configuration which requires three Wilkinson power dividers.

2.2.6.5 Edge-Coupled Feed Probe Design for RHCP Patch Antenna

In this type of feeding scheme the two sides of a square-patch antenna are each connected to separate feed microstrip feed lines, which are in turn connected to the output ports of a branch-line coupler as shown in Figure 2.23(a). This ensures that they will have equal amplitudes but with phases that differ by 90° for generating RHCP from one of the two output ports of the coupler. The other port of the coupler is terminated in a matched load. Impedance matching networks can also be inserted between the antenna and the feed lines connecting the antenna to the branch-line coupler to improve the impedance bandwidth. This type of feed network provides several advantages. The patch antenna, the polarizing feed, and impedance-matching network are coplanar and on the same substrate. This makes it simpler and less expensive to build and allows for an easier integration of other components such as bandpass filters or an LNA directly into the antenna circuit board; it is especially conducive for multilayered circuit boards often used in receivers. The disadvantage is that the antenna has a larger lateral dimension, making it unsuitable for many applications such as an adaptive antenna array, where a compact size antenna is essential because of the restriction on lateral dimensions. Another disadvantage is spurious radiation from the feed lines especially if there is a mismatch in impedance at the junction between the feed line and the patch antenna.

2.2.6.6 Proximity Coupled Feed Design for RHCP GPS Patch Antenna

A single proximity coupled feed for a RHCP patch antenna that operates at 1.575 GHz, the GPS L1 center band frequency, has been designed by Iwasaki and others [42] and is shown in Figure 2.23(b). This feed concept does not need an external hybrid; the advantage it provides is similar to that of single direct feed contact probe described earlier. This scheme also provides a choice of two dielectric
substrates—one for the patch and the other for the feed line so that the performance can be optimized. The rectangular-shaped radiating patch antenna is proximity-coupled to the feed microstrip transmission line that is offset from the center of the patch; there is no electrical contact between the feed line and the patch antenna. The four design parameters available with this design for obtaining both a good impedance match as well as axial ratio is the aspect ratio of the patch \( \frac{L_p}{W_p} \), where \( L_p \) and \( W_p \) are the length and width, respectively, of the rectangular patch; the offset length \( L_0 \) from the center of the rectangular patch to the center line of the microstrip line and \( S \); the distance from the edge of the patch antenna; and the edge of the microstrip line. By adjusting these parameters it is possible to excite two orthogonal modes of equal amplitude and with a phase difference of 90° at the selected design frequency. To obtain RHCP requires \( L_p < W_p \) and \( -\frac{W_p}{2} < L_0 < 0 \). The antenna model to demonstrate this concept was built on a copper-clad substrate with a dielectric constant of 2.6; unfortunately neither the thickness \( t \) of the patch antenna substrate nor the thickness \( h \) of the microstrip line substrate are provided in [42]. \( L_p = 56 \) mm and \( W_p = 58 \) mm for an aspect ratio \( \frac{L_p}{W_p} = 0.966 \); the width \( W_S \) of the microstrip feed line was 4 mm and \( L_0 = 14.4 \) mm and \( S = 13.8 \) mm. A 0.3-dB axial ratio at boresight and a 2-dB axial ratio over 60° angular range around boresight were measured. The bandwidth of the 2-dB axial ratio was 0.55% and the bandwidth for a VSWR< 2 was 3.5% at a center frequency of 1.575 GHz. The ground plane below the microstrip feed line prevents back radiation and yields a good FBR with this feed design. Spurious radiation from the feed line is also avoided. The resonant frequency of the antenna can be tuned by changing the dimensions of the patch antenna and also be fine-tuned further by changing the location of the open-circuit termination at the end of the microstrip feed line. If the patch substrate is made thinner a slight improvement in gain is observed and a thicker substrate produces the opposite effects. The operating frequency is increased when a thinner feed substrate is used. The disadvantage of this feed method is that accurate alignment is needed between the patch and the feed line and the axial ratio bandwidth is relatively narrow that is typical of a single-feed design.

### 2.2.7 Circular RHCP Microstrip Antenna

A picture of a dual-band circular microstrip patch antenna for GPS applications was shown earlier in Figure 2.5(b). Figure 2.28 shows the schematic diagram of a single-layer circular microstrip antenna fed by two feed probes; a cylindrical coordinate system \((\rho, \phi, z)\) represents the antenna geometry. The microstrip antenna consists of circular metal disc of radius \( a \) located at the center of a dielectric substrate of thickness \( h \), a dielectric constant \( \varepsilon_r \), and loss tangent \( \tan \delta \). The dielectric substrate is backed by a conducting ground plane. As in the square-shaped patch antenna, RHCP in the circular patch can be generated by using two direct-contact coaxial probes located at \( P_2(\rho_0, \phi) \) and at \( P_2(\rho_0, \phi + \alpha) \) The two probes are placed along orthogonal radii so \( \alpha = 90° \) and the radial distance \( \rho_0 \) of the two probes from the center of the patch antenna are adjusted so that their input impedance is 50 ohms and that they provide a good match to the receiver. The coaxial feeds are connected to a quadrature hybrid, which provides equal amplitudes but phases that differ by 90° to the two probes. The electric field \( E_Z \) of the circular microstrip antenna generated by a probe is given in the cylindrical coordinate system \((\rho, \phi, z)\) by:
where $J_n$ is the Bessel function of the first kind of order $n$, $k = k_0 \sqrt{\varepsilon_r}$ is the propagation constant in the dielectric substrate with relative dielectric constant $\varepsilon_r$; $k_0 = \frac{2\pi}{\lambda_0}$ and $\lambda_0 = \frac{c}{f}$ where $f$ is the frequency in cycles per sec, and $c$ is the velocity of light in free space. $E_0$ is the electric field at the edge of the patch across the gap.

2.2.7.1 Resonance Frequency and Radius of a Single-Layer Circular Microstrip Antenna

For each modal configuration of the circular patch antenna, a radius can be found that results in resonance when the derivative of the Bessel function of order $n$ is zero [23, pp. 73–85; 21, pp. 317–356].

$$J_n'(k_0 \sqrt{\varepsilon_r}a) = 0$$  \hspace{1cm} (2.32)

If $\chi_m$ is the $m$th zero of $J_n'(k_0 \sqrt{\varepsilon_r}a)$, the derivative of the Bessel function of order $n$. The resonance in the antenna occurs when $k_0 \sqrt{\varepsilon_r}a = \chi_m$, $n = 0, 1, 2, \ldots$, and $m = 1, 2, 3, \ldots$. In the $TM_{mn}$ mode of the circular microstrip antenna the index $n$ represents the angular mode related to $\theta$ and the index $m$ represents the radial mode.

\[ E_z = E_0 J_n \left( k_0 \sqrt{\varepsilon_r} \rho \right) \cos (n \phi) \]  \hspace{1cm} (2.31)
related to the radial coordinate. The mode number \( n \) corresponds to the number of sign reversals in \( \pi \) radians in \( \phi \). Value of \( \chi_{nm} \) for the first few \( TM_{nm} \) modes, each of which has its own unique radiation pattern, is given in Table 2.3.

The lowest-order mode in the circular patch antenna is the \( TM_{11} \) mode, with \( n = 1 \) and \( m = 1 \) is important for GNSS. This is a bipolar mode with the electric field concentrated at each end of the antenna, has the smallest radius or the lowest resonance frequency for a circular patch antenna, and is analogous to the lowest-order mode \( TM_{10} \) in the square-patch antenna. This is the primary mode of interest of the circular patch for GNSS; \( \chi_{11} = 1.841 \) for this mode.

The resonant frequency \( f_{nm} \) for the \( TM_{nm} \) mode of a circular microstrip antenna is

\[
f_{nm} = \frac{\chi_{nm} c}{2\pi a_{eff} \sqrt{\varepsilon_r}}
\]

where \( a_{eff} \) is the effective radius of the patch antenna and is given

\[
a_{eff} = a \left[ 1 + \left( \frac{2b}{\pi a \varepsilon_r} \right) \left[ \ln \left( \frac{\pi a}{2b} \right) + 1.7726 \right] \right]^{1/2}
\]

where \( a \) is the physical radius of the circular patch antenna.

The above two equations may now be combined to produce for the radius \( a \) for the \( TM_{11} \) mode

\[
a = \frac{1.8411 c}{2\pi \sqrt{\varepsilon_r}} \left[ 1 + \left( \frac{2b}{\pi a \varepsilon_r} \right) \left[ \ln \left( \frac{\pi a}{2f_{11}b} \right) + 1.7726 \right] \right]^{1/2}
\]

In the above equation \( f_{11} \) is the resonant frequency of the \( TM_{11} \) mode of the circular patch. The above equation is of the form \( a = f(a) \) and can be solved by using an iterative process to determine the radius of the patch antenna needed to obtain resonance for the \( TM_{11} \) mode at a resonance frequency \( f_{11} \). The initial approximation for the radius can be selected to be \( a_0 = \frac{1.8411 c}{2\pi f_{11} \sqrt{\varepsilon_r}} \). The initial value is placed on the right-hand side of (2.35) to produce the first iteration value of the radius designated as \( a_1 \). In the second iteration \( a_1 \) is placed on the right-hand side of the equation to obtain second iterative value designated \( a_2 \). In the third iteration \( a_2 \) is now substituted on the right-hand side to obtain a fourth value. This process is continued until it converges to a stable solution; generally 5 to 6 iterations are sufficient to produce the final result for the radius.

<table>
<thead>
<tr>
<th>Mode</th>
<th>( TM_{11} )</th>
<th>( TM_{21} )</th>
<th>( TM_{31} )</th>
<th>( TM_{41} )</th>
<th>( TM_{51} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \chi_{nm} )</td>
<td>1.841</td>
<td>3.054</td>
<td>4.201</td>
<td>5.317</td>
<td>6.415</td>
</tr>
</tbody>
</table>
2.2.7.2 Radiation Pattern, Feed Probe Location, and Gain of a Circular Microstrip Patch GNSS Antenna

The radiation pattern of a circular patch antenna for the $TM_{11}$ mode can be obtained from the cavity model and is given by [21]:

$$E_\theta = -jV \frac{ak_0}{r} \frac{e^{-ik_0r}}{r} J_1(k_0a \sin \theta) \cos(\varphi)$$  \hspace{1cm} (2.36)

$$E_\varphi = jV \frac{ak_0}{r} \frac{e^{-ik_0r}}{r} J_1(k_0a \sin \theta) \cos(\theta) \sin(\varphi)$$  \hspace{1cm} (2.37)

An RHCP-radiated field can be obtained by combining the outputs of the two feed probes located on the antenna at $P_1(\rho_f, \phi_f)$ and $P_2(\rho_f, \phi_f + \alpha)$ with equal amplitudes but with their phases differing by $90^\circ$ when the offset angle between the two feed probes $\alpha = 90^\circ$. The $\theta$ and $\varphi$ components of the total field $E_T$ can now be expressed as:

$$E_{T\theta} = E_{i\theta}(\varphi, \theta) - jE_{2\theta}(\varphi + 90^\circ, \theta)$$  \hspace{1cm} (2.38)

$$E_{T\varphi} = E_{i\varphi}(\varphi, \theta) - jE_{2\varphi}(\varphi + 90^\circ, \theta)$$  \hspace{1cm} (2.39)

In the above equations the subscripts 1 and 2 indicate the field contributions $E_1$ and $E_2$ made by the feed probes 1 and 2, respectively, to the total field $E_T$ along $\hat{\theta}$ and $\hat{\varphi}$. Although only two coaxial feed probes are considered in this discussion, some GNSS antenna designs use up to four probes to improve the circularly symmetry of the RHCP antenna pattern and to reduce phase center variations as shown in Figure 2.5(b). The locations of the additional two feed probes are selected to be diametrically across from the original two feed probes. However, using an increased number of probes will increase the complexity of the feed network and will reduce antenna gain because of the resulting increase in losses. The radial distance $\rho_0$ of the two probes on the patch antenna must be selected to provide an input impedance of 50 ohms to the input of the GPS receiver. The input resistance $R_{in}$ of the feed probe can be calculated from the equations below:

$$R_{in} = R_i \frac{j^2(k_{11}\rho_0)}{j^2(k_{11}a)}$$  \hspace{1cm} (2.40)

where $k_{11}a = 1.84118$

$$R_i = \frac{1}{G_r} = \frac{\lambda_0^2 \eta_0}{\pi^2 a^2} \left[ \frac{1}{4} - \frac{8(k_0a)^2}{15} + \frac{11(k_0a)^4}{105} \right]$$  \hspace{1cm} (2.41)
where:

\[ \lambda_0 = \text{wavelength} \]
\[ \eta_0 = 120 \pi = \text{free space impedance} \]
\[ a = \text{radius of the patch antenna for the } TM_{11} \text{ mode} \]
\[ R_r = \text{radiation resistance of the antenna} \]
\[ G_r = \text{the resonant radiation conductance of the circular patch antenna fed from the edge.} \]

\( D \) is the directivity of the antenna and the gain \( G \) of the antenna are given by the following equations:

\[
D = \frac{(k_0 a)^2}{120 G_r} \tag{2.42}
\]

\[
G_r = e_r D \tag{2.43}
\]

The radiation efficiency \( e_r \) is the ratio of the radiated power to the input power and is determined by the conductor losses in the patch antenna, the dielectric losses in the dielectric substrate, and the surface wave losses. The value of \( e_r \) is within the range 0 to 1. Hence, the gain is always less than the directivity.

Figure 2.29 shows the measured radiation pattern at 1.5754 and 1.2276 GHz of a stacked, dual-band, circularly shaped patch antenna designed for operation in the GPS L1 and L2 bands using the HFSS code. The RHCP is obtained by exciting both patch antennas with four coaxial probes. The substrate used for this circular patch antenna is TMM13i with a dielectric constant of 12.78 and a loss tangent of 0.0027. The substrate thickness for the top and bottom patches was each 0.2" and the diameter of the dielectric substrate was truncated. The radius of the top L1 patch was 1.138" inches and the radius of the bottom L2 patch was 1.285".

**Figure 2.29** Measured elevation plane radiation pattern of GPS dual-band stacked circular microstrip antenna at (a) 1.5754 GHz and (b) 1.2276 GHz. Azimuth angle = 0 degrees.
2.2.7.3 Bandwidth of a Circular Microstrip Patch Antenna

The percentage bandwidth $BW$ of a circular patch antenna for an input VSWR of $S:1$ is given by:

$$BW = \frac{100}{Q_T \sqrt{S}} \%$$  \hspace{1cm} (2.44)

where $Q_T$ is the total quality factor [21, pp. 349–350]. For a circular patch antenna operating in the dominant $TM_{11}$ mode and with a patch antenna made from copper

$$Q_T = \left[ \frac{2.09 \times 10^{-4}}{\left( \frac{b}{\sqrt{f}} \right)^2 + \tan \delta + bf (ak)} \right]^{-1}$$  \hspace{1cm} (2.45)

where $b = \text{height of the dielectric substrate in cms}$, $f$ is the frequency in GHz, $\delta$ is the loss tangent of the dielectric substrate, and $k$ is

$$I_1 = \pi \left[ \{A\}^2 + \cos^2(\theta) \{B\}^2 \right] \sin \theta d\theta$$  \hspace{1cm} (2.46)

where [21, 23]

$$A = J_2(k_0 a \sin \theta) - J_0(k_0 a \sin \theta) \quad \text{and} \quad B = J_2(k_0 a \sin \theta) + J_0(k_0 a \sin \theta)$$

The above equations show that the bandwidth of a circular patch antenna for a selected resonance frequency can be increased by choosing a thicker dielectric substrate and also by lowering the dielectric constant of the substrate material. However, lowering the dielectric constant of the substrate will increase the diameter of the circular patch antenna.

2.2.8 Annular Ring RHCP GNSS Microstrip Antenna

An annular ring microstrip antenna is formed by removing a concentric circle of the conducting surface at the center of a circular microstrip patch antenna [22, pp. 169–179]. Modified versions of this antenna have been used for several different GNSS applications including multipath suppression. A picture of a typical annular ring antenna is shown in Figure 2.30(b). A cross-sectional sketch of the annular ring patch antenna in Figure 2.30(a) shows an outer radius of $b$ and an inner radius $a$. The resonant frequency $f_{nm}$ of the annular ring patch antenna for the $TM_{nm}$ mode can be determined from

$$f_{nm} = \frac{(k_{nm}) c}{2\pi \sqrt{\varepsilon_r}}$$  \hspace{1cm} (2.47)
where $k_{nm}$ are the roots of the characteristic equation of the form

$$J_n'(CX_{nm})Y'_n(x_{nm}) - J_n'(X_{nm})Y'_n(CX_{nm}) = 0$$

(2.48)

In the above equation $a$ and $C = \frac{b}{a}$ and $J_n(x)$ and $Y_n(x)$ are Bessel functions of the first and second kind of order $n$, respectively, and the prime denotes the derivatives with respect to $x$. For the case $C = 2$ and for the $TM_{11}$ mode with $n = 1$ and $m = 1$, the lowest order mode of the annular ring patch antenna, the inner and outer radii, can be determined from $x_{nm} = 0.6773$. An annular ring patch antenna excited in the $TM_{11}$ mode is shown in Figure 2.30(b). RHCP is obtained by exciting the antenna shown in this picture by four symmetrically located direct-contact probes excited with equal amplitudes and $0^\circ$, $-90^\circ$, $-180^\circ$, and $-270^\circ$ in phase, respectively, by using suitable hybrid couplers. The four probes need to be located at appropriate positions inside the patch antenna so as to provide an input impedance of 50 ohms. The impedance is at its highest value near the outer edge of the annular patch and lowest at its inner edge. The correct location of the probe is determined by using one of the computer design codes mentioned earlier or through analytical formulas [21, pp. 178]. Dual-band GNSS antennas can also be built by stacking a higher-frequency patch on top of a lower-frequency patch and exciting both patches with a common set of direct-contact probes.

Additional modifications to the radiation pattern can also be achieved by shorting either the inner and outer surfaces of the annular patch antenna and adjusting the inner and outer radii $a$ and $b$ so as to achieve resonance at the desired GPS frequencies. Dual-band annular ring patch antennas can also be obtained by stacking. One such useful modification for limiting GPS multipath effects is by using a reduced surface wave antenna described below.
### 2.2.8.1 Shorted Annular Ring Microstrip Antenna for GPS Multipath Limitation

As discussed in Section 2.2.4.4, the $TM_0$ surface wave mode in a microstrip antenna has no cutoff frequency. Lateral waves are also launched at the surface of the ground plane [25]. Hence they can generate antenna backlobes that can increase multipath effects when both the surface wave and lateral wave that are launched by the microstrip antenna are diffracted from either the truncated edges of the dielectric substrate of the antenna or the ground plane on which the antenna is mounted (shown in Figure 2.6). Multipath arises when the signals transmitted by the GPS satellites arrive at the receiving antenna using two or more different paths—one from the main beam and the other from extraneous reflected signals picked up from the antenna backlobes—resulting in amplitude and phase differences caused by their interference. These reflected signals often impinge on the antenna at low-elevation angles. The range distance measurement from the satellite to the user is corrupted by such multipath interference and the error depends on the time delay between the direct and reflected signals and the strength of the reflected signal relative to the direct signal. Multipath effects and their impact on GNSS measurement accuracy are described in greater detail in Chapter 5 and 7.

Various types of multipath limiting antennas have been designed for GNSS applications. Unlike different types of choke ring designs that rely on complex corrugated ground planes to suppress surface waves, the shorted annular ring (SAR) antenna accomplishes the same result of reducing the horizon and backlobe radiation by using an antenna design that is simpler, lighter, and less expensive to build [19, 20]. In this design the inner circumferential edge of the annular ring patch antenna is shorted to the ground plane and both the inner and outer radii of the antenna are adjusted to suppress the $TM_0$ surface wave as well as the lateral wave propagation. A schematic diagram of the annular microstrip ring patch antenna with its inner circumferential edge shorted to the ground plane is shown in Figure 2.31. The $TM_0$ surface wave and the lateral wave are suppressed when the outer radius $b$ of the annular ring patch is adjusted to satisfy the condition:

![Figure 2.31](image)

**Figure 2.31** SAR microstrip antenna for reducing multipath.
where \( k_\circ = \frac{2\pi}{\lambda_0} \) and \( k_{TM0} \) wave propagation constant of the \( TM_0 \) surface wave mode in the dielectric substrate of the patch antenna, and \( \lambda_0 \) is the wavelength in free space. If the dielectric substrate is electrically thin (i.e., if \( h \ll \lambda_0 \)), then \( k_{TM0} \approx k_\circ \) and the condition specified in (2.49) is satisfied. Hence the outer radius \( b \) of the annular ring microstrip antenna that is necessary for achieving maximum suppression of both surface wave and lateral wave propagation is given by:

\[
b = \frac{1.8412}{k_\circ} = 0.29 \lambda_0 \quad (2.50)
\]

The above equation indicates that for a thin substrate the outer radius of the annular ring antenna necessary for suppressing both surface wave and lateral wave radiation is independent of the dielectric constant and is 0.29 times the wavelength. This means that the outer radius needs to be 2.191" at 1.57654 GHz and 2.819" at 1.2276 GHz, the center frequencies of the GPS L1 and L2 band, respectively. To make the annular ring antenna resonate at the desired GPS frequency a short-circuited inner boundary needs to be incorporated into the design. In some designs the short circuit to the ground plane is obtained by a concentric array of closely spaced vias at the inner surface of the annular ring; in other designs a concentric circumferential metal ring is soldered to the top and bottom inner surface of the annular ring patch antenna. To obtain resonance the radius \( a \) of the short-circuited inner boundary should satisfy the transverse resonance equation given by:

\[
\frac{\left[ J_1(k,a) \right]}{\left[ Y_1(k,a) \right]} = \frac{\left[ J'_1(k,b) \right]}{\left[ Y'_1(k,b) \right]} \quad (2.51)
\]

where \( k_1 = k_\circ \sqrt{\varepsilon_r} \) and \( \varepsilon_r \) is the dielectric constant of the substrate.

Two types of dual-band SAR antennas for multipath mitigation have also been proposed for the GPS L1 and L2 bands. The first design uses stacked patch antennas with elliptically shaped patches for achieving dual-band performance; both top and bottom patches are excited by a common set of direct-contact probes [43]. In the second design resonance in the GPS L2 band is obtained by using a second inverted shorted annular ring (ISAR) microstrip antenna where the outer radius is shorted to the ground plane instead of the inner radius [43]. This “shorting” is achieved by using a concentric array of closely spaced vias similar to that used for the inner annular ring. The ISAR antenna is made to resonate in the L2 band by adjusting the outer and inner radii of the outer ring, \( a_I \) and \( b_I \), respectively, to satisfy the transverse resonance equation [43]:

\[
\frac{\left[ J_1(k,a_I) \right]}{\left[ Y_1(k,a_I) \right]} = \frac{\left[ J'_1(k,b_I) \right]}{\left[ Y'_1(k,b_I) \right]} \quad (2.52)
\]
Figure 2.32(a) shows a schematic diagram of the second design concept by combining a SAR antenna operating in the \( L_1 \) band with an ISAR antenna operating in the \( L_2 \) band. Dimensional and other design details are discussed in a paper by Basilio et al. [43]. Figure 2.32(b) shows the measured radiation pattern of a dual-band multipath limiting antenna. The first dual-band antenna design is more compact and is also compatible for use with a dual-band GPS receiver since it uses a common set of feed probes for both frequency bands. The second design is much wider in circumference and is also less compatible for use with a dual-band receiver unless used with a diplexer.

### 2.2.9 Mutual Coupling Effects in GNSS Microstrip Antenna

Mutual coupling is always present between elements in a microstrip antenna array and is a function of the edge separation between adjacent elements, the dielectric constant the thickness of the substrate [23, pp. 163–174]. In microstrip antenna elements mutual coupling can occur from both space wave as well as surface wave radiation, especially through the \( TM_0 \) surface wave, which has no cutoff even with thin substrates. Mutual coupling effects are generally negligible when the spacing between elements is a half-wavelength or greater. As the spacing between the elements decreases, the mutual coupling increases rapidly and can have a significant negative impact on almost all of the important parameters needed for optimum GNSS performance such impedance mismatch to the feed network, degradation in RHCP gain and axial ratio, pattern asymmetry with elevation and azimuth angle, and increased phase and group delay. Mutual coupling effects can be significant in small and finite size array antennas placed on aircraft where the antenna elements are packed much closer than a half-wavelength apart in the interests of conserving space available for deploying antennas [44]. Previous studies have shown that mutual coupling between antenna elements can affect the channel mismatch in GPS antenna arrays and can affect the interference cancellation ratio (ICR), and signal-to-interference-plus-noise ratio (SINR). Mutual coupling effects in microstrip

![Figure 2.32](image-url)
antennas vary differently in the E and H planes and affect the axial ratios in RHCP elements used in GNSS. For close interelement spacing, mutual coupling in the H plane are larger than in the E plane; the H-plane coupling and can be as large as -10 dB when the edge spacing decreases to 0.15 wavelengths. The differences in the measured mutual coupling between two linearly polarized square-shaped microstrip antennas in the E and H plane as a function of their relative edge spacing distance have been measured by Jedlicka [45] and are shown in Figure 2.33.

This imbalance in mutual coupling in the E and H planes can affect the amplitude and phase balance between the fields generated by the two orthogonally located feed probes in a patch antenna that generate the RHCP [44]. In an isolated antenna element, the two probes have equal amplitude with a phase difference of 90°. The E- and H-plane asymmetry in mutual coupling alters this phase and amplitude balance in an antenna element placed inside an array and increases the cross-polarization level depending on the severity of the coupling between adjacent elements. Mutual coupling between the two separate feed probes also excites higher-order modes, especially the $TM_{12}$ mode in a square-patch antenna that can cause the radiation pattern to degrade [46]. This is particularly true on a patch antenna fed by two feed probes because of the inherent asymmetry in excitation. The problem can be remedied to some extent by using symmetrically placed four-feed probes as mentioned in the next section. In a closed packed array of microstrip antenna elements the dielectric substrate below each patch antenna is often truncated to minimize direct surface wave coupling between adjacent elements.

Ngai and Blejer [47] have calculated the mutual coupling for a number of different element configurations used in GPS microstrip antenna arrays. Figure 2.34 shows the mutual coupling between two identical square-shape RHCP antennas at two different relative orientations for substrates with three different dielectric constants. Note that mutual coupling for the diagonal configuration is greater than for the horizontal or equivalent vertical configuration. The diagonal configuration

![Figure 2.33 Measured mutual coupling between two microstrip antennas in the E and H planes. (From [45]. ©1981 IEEE.)(From [45]. ©1981 IEEE.)](image-url)
is the geometry that is prevalent in the mutual coupling that exists between the central element (generally the key reference element of an adaptive array) and the outer auxiliary elements of a square array. Ngai and Blejer [47] have also done extensive calculations of mutual coupling effects in arrays of microstrip antennas containing from 3 to 9 antenna elements enclosed in a $7'' \times 7''$ square aperture. To keep the element size small the dielectric constant of the substrate used in these simulations was 10.2. Calculations have been made for both the GPS L1 and L2 frequency bands. The results for mutual coupling between elements in GPS antenna arrays measured by Nagai and Blejer are shown in Figure 2.35.

2.2.10 Phase and Group Delays in GNSS Microstrip Antennas

Phase and group delays introduced by microstrip antennas are a potential source of error in high-precision GNSS; there have been several investigations to estimate their magnitude [34, 38, 48, 49]. These errors can vary with both the elevation and the azimuth angle and thus can be different for each satellite signal that is received, leading to positional errors. Group delay errors tend to be much larger than phase delay errors in microstrip antennas due to the high Q factor of the microwave cavity formed by the dielectric substrate sandwiched between the top metallic patch and the ground plane. This restricts the antenna bandwidth and is particularly true for substrates with high dielectric constants such as ceramics. Carrier phase-based methods such as precise point positioning (PPP) and real-time kinematic (RTK) are particularly susceptible to antenna-induced errors; they can be compensated by calibrating the PCO and PCV of the antenna.

Dong et al. [34] have investigated the phase and group delay as a function of elevation and azimuth angles introduced by GPS microstrip antennas with four different types of direct feed probe configurations. The simulation was made in the
GPS L₁ frequency band. The four patch antenna designs considered in this study were:

1. A square-shaped patch antenna with a pair of orthogonal probes excited with a relative phase of 0° and −90° (discussed in Section 3.6.1).
2. A nearly square patch antenna with a single diagonal probe (discussed in Section 3.6.3).
3. A square-shaped patch antenna consisting of two pairs of symmetric orthogonal probes for a total of four probes; these are excited with relative phase of 0°, −90°, −180°, and −270°, respectively (discussed in Section 3.6.2).
4. A nearly square patch antenna with two symmetric diagonal probes; these are excited with relative phase of 0° and 180°. This type of feeding scheme was not discussed earlier but is shown in Figure 1(d) in [34].

The coaxial probes were SMA connectors with a probe radius of 0.635 mm. The phase and group delays were analyzed using the full wave electromagnetic simulator HFSS by Ansoft. A cavity model was also used in the analysis but provided less accurate predictions since it is unable to fully take into account effects from higher-order modes. The dielectric constant of the substrate used in most of the cases shown in the paper was 2.94, although in one case simulations were also made for orthogonal feed patch antennas with substrates with dielectric constants...
of 1.0, 2.94, and 10.2. The results of this study show that the group delay for the diagonal feed patches are much higher than for the orthogonal feed patches. For both the orthogonal- and diagonal-feed configurations the symmetric version with four probes had smaller group delays than the corresponding asymmetric two-feed orthogonal or the one-feed diagonal configuration. This is because the symmetric configuration is able to suppress the higher-order modes from being excited in the patch antenna. The symmetric, orthogonal probe feed antenna with four coaxial probes, discussed in Section 2.3.6.2, showed the best symmetry for group delay variation in the azimuth plane; it does this by suppressing the excitation of the higher-order modes other than the dominant (1,0) and (0,1) modes. In contrast, the diagonal-feed patch antennas showed the greatest variation in both elevation and azimuth angles with in some cases group delays of the order of 4 ns; large asymmetry in the azimuth plane is also noticed with the diagonal feeds.

The primary contributor to group delay is the narrow bandwidth of the patch antennas. The group delay increases with the dielectric constant of the substrate due to the narrower bandwidth. The phase delay error is not large for all feed configurations and is of the order of millimeters for a typical patch antenna. The primary contributor to the phase delay is probe reactance of the coaxial feed probe. Murphy et al. [48] have also investigated group delay from microstrip patch antennas via modeling using the FEKO, Ver. 5.3 electromagnetic simulation code. The patch was excited by two-feed probes orthogonally located relative to each other within the patch; the patch conductor was 0.0287 m square on a finite-size dielectric substrate that was 0.0445 m square and 0.00254 m thick. The substrate was Roger’s 6010 LM with a dielectric constant of 10.2 and a loss tangent of 0.0023. The ground plane was 2-feet square with rolled edges to reduce edge diffraction effects. The FEKO model of this patch antenna was used to compute the phase and amplitude phase response of the antenna pattern in the GPS L1 band. The variation in the group delay as a function of elevation and azimuth in 1° increments above the horizon was calculated. These computations were repeated for multiple frequencies about the center frequency $f_c = 1575.42$ MHz in the GPS L1 band. A completed frequency domain transfer function was obtained for each angle of arrival. The differential group delay at $f_c$ was obtained based on the first difference of phase versus frequency as predicted by the model. The variation of the differential group delay was calculated as a function of the azimuth and elevation angles of the satellite. The maximum group delay variation is less than 1.5 ns and appears to be in general agreement with the results shown by Van Dierendonck et al. [49] based on measurements made on two commercial GPS microstrip antennas used in aircraft navigation.

Wirola et al. [38] have investigated the antenna phase response and their effects on ambiguity resolution for three antennas in the GPS L1 band. Two of these antennas were custom-made microstrip patch antennas with a single diagonal-feed probe used in Bluetooth- (BT) assisted GPS (AGPS) receivers, referred to as BAG. These are 25 x 25 mm square-patch antennas with one corner truncated. The substrate is a ceramic with a dielectric constant $\varepsilon_r = 20$ and has a $Q = 5000$. The performance of these two patch antennas was compared against that of a third antenna: Trimble Bullet™ III. The full 3-D complex RHCP far-field response as a function of elevation and azimuth for all three antennas were measured using the Satimo SG128 3-D measurement system. The $S_{11}$ parameter of the antennas is also provided in the
The results show a clear contrast in response between the inexpensive BAG ceramic patch antennas and the Trimble Bullet antenna. Details on the design of the Trimble Bullet antenna are, unfortunately, not provided in the paper.

The Bullet antenna shows a highly symmetric amplitude and phase response and has less than a 4° variation in phase in any given direction, which corresponds to an error of less than 1.6 mm. The standard deviation of the phase response from a best fit sphere is only 0.7°. In contrast the two BAG antennas that were tested showed maximum variations of 70° and 49° in their phase response with standard deviations around the best fit sphere of 8° and 6°, respectively. Their amplitude responses showed similar asymmetry with the observation angles. The greatest variation was observed at the lower-elevation angles. When the elevation masking angle was raised to 30° or 45°, the maximum phase response variation decreases to 32° and 19°, respectively. In addition the response for the two antennas showed differing phase responses so that a single PCV table for this type of antenna is not feasible. The large phase errors observed with the two BAG antennas did not appear to significantly affect ambiguity resolution according to Wirola [38]. These results seem to support observation by Dong et al. [34]. Diagonally fed, single-probe patch antennas create large phase response errors that can affect precision GPS measurements due to the excitation of higher-order modes. More uniform phase and amplitude response as a function of observation angle can be obtained by replacing these with patch antennas fed with four feed probes symmetrically arranged around the center of the square patch. However, the use of four probes will reduce the overall gain of the antenna due to the higher feed losses from the more complex feed network that is needed.

### 2.2.11 Advantages and Disadvantages of Microstrip Antennas for GNSS Applications

The popularity of microstrip patch antennas for GNSS stems from the numerous advantages discussed in the previous sections and is summarized below.

**Advantages**

- They can be made to have a very low profile and hence are especially suited for avionics since they can easily meet the ARINC 743 requirement for GPS avionics antennas. This requirement stipulates that the maximum antenna height be no greater than 0.73" so as not to increase aerodynamic drag. Figure 2.3 shows the ARINC 743A antenna footprint for such airborne antennas.
- By using substrates with high dielectric constants listed in Table 2.1, the size of the antenna can be made very small, which makes microstrip antennas ideally suited for densely packed antenna arrays. Very compact ceramic loaded antennas are also common in GPS handsets since they can be unobtrusively mounted inside the receiver casing; their design is described in Chapter 3.
- High dielectric constant substrates enhance the end-fire surface wave radiation from the patch antenna. This improves directivity of single element GNSS antennas at very low-elevation angles near the horizon. The antenna is
therefore able to provide better gain at the lower-elevation angles to acquire
GNSS satellites that are critical for improving PDOP. Surface wave radia-
tion from microstrip antennas can also be suppressed by using SAR patch
antennas. Although this greatly reduces the gain at low elevation angles and
affects satellite availability, it is nevertheless useful as a technique for limiting
multipath for certain applications.

- Dual-band antennas covering the GPS \( L_1 \) and \( L_2 \) bands or the \( L_1 \) and \( L_5 \)
bands needed for correcting propagation errors caused by the ionosphere
can be made by either stacking patch antennas on top of one another or by
using parasitic coupling of adjacent patches. Recently, triple-band microstrip
antennas covering the modernized GPS (\( L_1, L_2, \) and \( L_5 \)) or the Galileo system
(\( E_5a, E_5b, E_6 \)) have been built using these design principles (discussed in
Chapter 3). By providing narrowband spot coverage at only the selected fre-
quency bands of operation the microstrip antenna also serves as a passband
filter to reduce out-of-band interference.

- RHCP can be easily achieved by using a variety of different feed structures
such as coaxial probes, aperture coupled slots, proximity coupled microstrip
lines, or edge coupled microstrip lines directly attached to the patch anten-
as. RHCP with a good axial ratio over only a narrow bandwidth is possible
by using a single feed probe coupled to a patch antenna whose external shape
is changed by either truncating the corner or by inserting indentations or tabs
on the side. This type of antenna feed does not need an external quadrature
hybrid and is therefore popular for compact, simple, and low-cost ceramic
antennas used in some handheld GPS receivers.

- Conformal mounting on surfaces of sleek aerodynamic shapes such as cones
and cylinders is possible by using thin flexible substrates making microstrip
antennas popular for small missiles or other projectiles used in many military
applications. Special flexible fabrics have recently been used for placing GPS
antennas directly on garments.

- Antennas can be very light in weight and can easily be integrated into micro-
wave integrated circuits used in GNSS receivers. Photolithography combined
with etching techniques lowers fabrication costs for high-volume production.

**Disadvantages**

Microstrip antennas also suffer from several disadvantages that provide a challenge
to the antenna designer when optimizing their performance for GNSS systems.

- They have a very narrow bandwidth requiring special design techniques to
achieve operation at two or more frequency GNSS bands such as stacking
one patch on top of another, or coupling to another parasitic patch tuned to
a different frequency, or by etching slots into the metallic patch.

- Dielectric substrates with high dielectric constants while reducing the size
of the antenna also decrease the bandwidth and increase surface wave ra-
diation at low-elevation angles near the horizon. While this provides some
advantages such as improved PDOP due to increased beam width, it also
causes increased diffraction from the edges of the ground plane on which the antenna is mounted. This results in backlobes that increase susceptibility to multipath and interference and also ripples in the antenna gain and phase at higher elevations due to interference with the direct radiation from the patch antenna. These effects are discussed in Chapter 5.

2.3 Quadrifilar Helix Antenna

The quadrifilar helix antenna (QHA), also known as a “volute,” is popular for many GNSS applications and is especially suitable for use in handheld receivers and mobile terminals. It is an array of four helically shaped resonant antenna elements wrapped around a cylinder at the desired pitch angle and driven in phase quadrature to produce a broad, RHCP beam. The QHA can be fed from its terminals at either the top end or the bottom end with the other end either shorted or left open and can therefore generate both forward and backward wave radiation patterns with different beamwidths and shapes, adding to its versatility for various GNSS applications. Its size can also be controlled by using dielectric substrates and printed versions can be made to reduce cost. Two types of QHAs are used in GNSS: self-phasing QHAs shown in Figure 2.36(a) and externally phased QHAs shown in Figure 2.36(b).

2.3.1 Self-Phasing QHA

The half-turn self-phasing QHA shown in Figure 2.36(a) consists of four helical elements connected to a split sheath coaxial balun at the top and with the bottom ends electrically shorted to the outer sheath of the coaxial line [50–53]. The QHA

![Figure 2.36](image.png)

Figure 2.36  Diagram of (a) self-phasing QHA and (b) externally phased QHA. (Figure 2.36(a) Courtesy of S. Best.)
in this design is fed from the top and operates in the backfire mode since the main radiation is directed from the bottom shorted end of the antenna towards the feed section at the top. Quadrature phasing required for generating RHCP is obtained by adjusting the lengths of four elements (described in greater detail later). Figures 2.37(a) and (b) show the side and top view, respectively, of this QHA without the radome cover normally placed around the QHA. The measured principal RHCP and the cross-polarized LHCP patterns of this antenna are shown in Figure 2.37(c). They indicate a very broad antenna beam providing an almost hemispheric coverage with good visibility to all available satellites. One of the drawbacks of QHAs fed with conventional coaxial line balun is considerable backlobe radiation resulting in poor front-to-back ratios, which makes this antenna susceptible to multipath. However, more recent QHA designs use infinite microstrip balun feeds that allow significant reduction in backlobe radiation and are able to achieve front-to-back ratios of 20 dB. The measured return loss of this self-phased QHA is shown in Figure 2.37(d). This antenna is resonant at a frequency of 1.5756 GHz and indicates a narrow bandwidth typical of a highly resonant, single-band antenna structure.

2.3.1.1 Design Principles of a Self-Phasing Quadrifilar Helix Antenna

The self-phasing QHA consists of four helical arms, each with an $n/4$ turn where $n = 1, 2, 3, 4$, and so forth, which are equally spaced circumferentially and then wrapped with a constant pitch around a small-diameter cylindrical dielectric core to form a helix. To make the antenna resonate, the four arms must be equal to $m\lambda/4$,
where \( \lambda \) = the wavelength and \( m = 1, 2, 3, 4 \ldots; \) the ends of the helices must be short-circuited when \( m \) = an even integer and open when \( m \) = an odd integer.

The QHA can be considered a combination of two orthogonal, uncoupled bifilar antennas fed in phase quadrature, where a bifilar is defined as a two-element helical antenna array. The top end of the four helical lines in self-phasing QHAs are connected to a coaxial balun, which is an acronym for a balanced to unbalanced transformer. The balun provides correct current paths between a balanced configuration such as the feeding coaxial cable and the unbalanced configuration such as the antenna. Baluns with different designs are used in QHAs [50, pp. 166–169, 54]. Figures 2.36(a) and 2.37(a) shows the geometry for a half-turn QHA fed by the coaxial feed line and a split-sheath type of balun located along the axis of the cylinder. This balun is designed such that the balanced outputs have equal amplitudes but are in antiphase (i.e., have a relative phase difference of \( 180^\circ \)).

Two adjacent elements of the quadrafilar are connected to the center conductor which is one-half of the balun. The other two elements of the quadrafilar are connected to the outer conductor, which is the other half of the balun. Therefore two of the elements are at a \( 0^\circ \) phase angle and the other two are at a \( 180^\circ \) phase angle. Adjacent helical elements at the same phase are trimmed to different lengths. One element is made shorter than the ideal resonant length to the shorting plane of the balun to make it capacitive at the desired frequency and is cut such that the current leads by \( 45^\circ \). The other element in this pair is cut longer than the correct resonant length to the shorting plane of the balun to make it inductive and causes the current to lag by \( 45^\circ \). The lengths of the other pair of adjacent elements, which are \( 180^\circ \) out of phase with the first pair, are adjusted the same way. The four elements of the QHA connected to the balun are now fed with equal amplitudes and with relative phases in the counter clockwise direction of \((0^\circ+45^\circ), (0^\circ-45^\circ), (180^\circ+45^\circ = -135^\circ), \) and \((180^\circ-45^\circ = -225^\circ)\). The \( \theta \) and \( \varphi \) components of the radiated electric field of the combined bifilars have the same amplitude but are in-phase quadrature to generate RHCP. The normalized input impedance of the longer, inductive bifilar is \( Z_L/Z_0 = 1+j \). Similarly the input impedance of the smaller, capacitive bifilar is \( Z_S/Z_0 = 1-j \). The normalized admittances of the two bifilars are given by \( y_L = 1/(1+j) \) and \( y_S = 1/(1-j) \). Since the two bifilars are fed in parallel, their combined admittance is the sum of \( y_L + y_S = 1 \), and the QHA is matched in impedance to the coaxial feed line [55]. To obtain an elevated omnidirectional antenna pattern, the sign of the relative phase difference between the helix elements and the direction of the helix winding occur in the opposite direction. The feeding direction of the quadrature phasing determines the end-fire direction of the antenna. A relative phase increment of \( +90^\circ \) in the clockwise direction between the individual helical elements produces an RHCP omnidirectional pattern if the helices are wound in a left-hand sense. Note that in a conventional end-fire helix, the helix has to be wound like a right-hand screw to obtain RHCP. Since the QHA operates in the backfire mode, with the main radiation directed towards the feed point and away from the shorted end, the helix has to be wound as a left-hand screw in addition to the correct phasing on the four helical wires [53, 54].

Since a CP signal changes its sense of polarization after a reflection, the wires of the QHA need to be wound in a left-hand sense so that upon reflection from either the ground plane in an externally fed QHA or the short circuit termination at the
bottom end in a self-phased QHA, the signal radiated from the top (feed) end of the antenna is RHCP.

2.3.1.2 Computational Design of a Self-Phased QHA

Kilgus [56] was able to provide approximate, closed-form analytical formulas for the radiation patterns of ¼-, ½-, and a l-turn QHA by approximating the QHA as a combination of a loop and a dipole. However, ensuring a more accurate design of the QHA for GNSS applications, which includes estimating both the amplitude and phase for the $\theta$ and $\varphi$ polarization components as well as the effects introduced by the central dielectric core or by the ground plane as in an externally phased QHA requires analysis using more advanced electromagnetic computational modeling methods. Method of moments (MOM) codes such as NEC-4 are useful for the analysis of QHAs that use a core made from a very low-dielectric constant such as foam or an air core. Accurate design of a QHA with a dielectric core with higher-dielectric constant than foam requires using commercial computational codes listed in Table 2.2, which are the same as those used for designing microstrip antennas.

Using these electromagnetic codes will require segmenting both the four helical and radial elements of the QHA into finite segments, which are fractions of a wavelength at the desired design frequency. Figure 2.38(a) shows the segmentation of a QHA for analysis using a MOM code. The geometrical equations for the QHA described below can be used to assist in this segmentation. The geometry of each of the QHA can be described in terms a helical coordinate system as shown in Figure 2.38(b).

The coordinate system shown in Figure 2.38 is defined by the following parameters: the pitch angle $\psi$ of the helix given by $\cot \psi = \frac{2\pi b}{P} = \tau b$ where $\tau = \frac{2\pi}{P}$, where $P$ is the pitch distance of the helix or the spacing between helical turn measured center to center, $b =$ the radius of the helix to the center line of the wire

![Figure 2.38](image-url)  
*Figure 2.38* Electromagnetic computational model of a self-phasing QHA. (Courtesy of S. R. Best.)
as shown in the figure, and \( a = \text{radius of the antenna wire.} \) The total axial length of the helix is \( L_{ax} = N \, P \), where \( N \) is the total number of turns (or a fraction of a turn) in the helix. Any point at an axial distance \( z \) on each of the four helical wires originating from the feed region at the points A1, A2, A3, and A4 at the top of the balun can be represented by the following four equations:

\[
A_1 = \hat{x}b \cos(\tau z) - \hat{y}b \sin(\tau z) + \hat{z}z \quad (2.53)
\]

\[
A_2 = -\hat{x}b \cos(\tau z) + \hat{y}b \sin(\tau z) + \hat{z}z \quad (2.54)
\]

\[
A_3 = \hat{x}b \sin(\tau z) + \hat{y}b \cos(\tau z) + \hat{z}z \quad (2.55)
\]

\[
A_4 = -\hat{x}b \sin(\tau z) - \hat{y}b \cos(\tau z) + \hat{z}z \quad (2.56)
\]

In the above equations, \( \hat{x}, \hat{y}, \) and \( \hat{z} \) represent the unit vectors along the three principal axes of the Cartesian coordinate system. Notice that helical wires originating from A1 and A2 form one pair of bifilars and those originating from A3 and A4 represent the orthogonal bifilar of the QHA. From the above equations it can be seen that for a half-turn QHA shown in Figure 2.38(b) the coordinates of the these four points at the feed plane \( z = 0 \) are:

\[
A_1 = \hat{x}b, \quad A_2 = -\hat{x}b, \quad A_3 = \hat{y}b, \quad \text{and} \quad A_4 = -\hat{y}b
\]

Similarly, at the other end of the QHA away from the feed point, where \( z = L_{ax} = -P/2 \) for the half-turn QHA, the four helical wires are “shorted” to the outer sheath of the feeding coaxial balun. This is the shorting plane of the balun and is used as the reference plane for adjusting the lengths of the helical arms to obtain the quadrature phasing. The coordinates of the corresponding points of the four helical wires at this shorting plane represented by \( A'_1, A'_2, A'_3, \) and \( A'_4 \) are given by:

\[
A'_1 = -b\hat{x} - \hat{z}(P/2), \quad A'_2 = b\hat{x} - \hat{z}(P/2), \quad A'_3 = -b\hat{y} - \hat{z}(P/2), \quad A'_4 = b\hat{y} - \hat{z}(P/2)
\]

Equations (2.53) through (2.56) can be applied to other QHAs with different fractional turns.

The elevation pattern or the cone aperture of the radiation pattern of the QHA can be controlled by adjusting the radius of the helix \( b \), the pitch distance of the \( P \), and the number of turns \( N \). These three parameters can be optimized for providing a near-ideal RHCP pattern covering the upper hemisphere. For GPS applications typical QHA designs use from one-quarter to a full turn to keep the length of the antenna small. The optimum number of turns for a QHA used in some GPS and GLONASS antenna systems is between 1.5 to 3 turns. Increasing the length of a QHA by increasing the number of turns does not have much impact on the gain, but improves the beam shape and attenuates the undesired backward antenna lobe in the downward direction, thereby improving its resistance to multipath from
ground reflections. Increasing the length of the antenna but maintaining the same number of turns reduces the front-to-back ratio; it will also increase the beamwidth and reduce the phase center variations at low-elevation angles. The diameter of the QHA plays an important role in determining the bandwidth with the sleeker style QHAs having a narrower bandwidth while those with more bandwidth require a larger diameter. The pitch angle, with the other parameters adjusted to be optimal, controls the beamwidth of the antenna. Higher pitch angles yield wider beamwidth at lower gain, whereas the lower pitch angles tend to yield narrower beamwidth at higher gain. The number of turns in itself is not a good indicator of the performance of the QHA.

Tranquilla and Best [57] by using MOM have provided detailed design information of the QHA with a foam (air) core specifically designed for GPS applications. They have considered both the amplitude and phase response of three different designs of the QHA summarized in Table 2.4. Of special importance is information they provide on the phase center stability of the QHA that strongly affects the use of carrier phase data that is important in differential GPS systems. The three antennas analyzed were one-half turn and one-half wavelength long measured along the curved helical segments including the radial section connecting to the central coaxial feed.

Figure 2.39(a) shows the computed RHCP antenna patterns for these three antenna designs. While all three QHAs provide good upper hemispheric gain coverage needed for acquiring visible GPS satellites, only one the shortest of these three with $L_{ax} = 0.2 \lambda$ has a suppressed backlobe that is need to mitigate ground multipath. Figure 2.40(b) shows the phase patterns of the linear polarized $\theta$ and $\phi$ components in a fixed observation plane; the phase data in this referenced paper [68] has been presented in linear polarization rather than in CP to maintain the strict definition of “antenna phase center”; the authors claim that phase center for CP is not a definable quantity although this is the polarization of interest for GNSS applications.

Figure 2.40(a) shows the computed phase center location for the $\phi$ and $\theta$ linear polarizations for the half-wavelength QHAs. The phase center is measured in terms of a length $L$ from the coordinate origin and an angle $\psi$ as measured from the axis $X$ of the QHA as shown in Figure 2.40(b). The improved phase stability

![Figure 2.39](image1.png)  
(a)  

![Figure 2.40](image2.png)  
(b)

**Figure 2.39** Calculated radiation pattern of a self-phased QHA. (From [68]. ©1990 IEEE.)
provided by these longer and slimmer QHAs is due to the reduced movement in the phase center at low-elevation angles. This is attributed to the reduced separation distances between helical arms when viewed from near broadside to the antenna, which minimizes the interference effects between the field contributions from the opposing arms of the QHA.

Hussain and Rengarajan [58] have conducted investigations using MOM techniques on a half-turn QHA with shaped ground planes specifically designed for GPS applications to examine their effects on the antenna backlobes in the radiation pattern as well as the impact on phase center stability caused by diffraction and mutual coupling from the edges of the ground plane. Both a flat ground plane $1\lambda \times 1\lambda$ and a bent ground plane with their edges bent downwards and away from the antenna at a $20^\circ$ inclination were considered. The ground plane was placed at various distances from the base of the antenna and its effects on the antenna pattern and phase center were examined.

QHAs designed for GPS applications have a narrow bandwidth of between 3% to 5% for a return loss of $-10$ dB VSWR = 2:1; this allows operation only over a single GNSS frequency band. As stated earlier, the natural bandwidth of the QHA is proportional to the diameter of the QHA; hence, the smaller the antenna the smaller its bandwidth. The return loss of a QHA used in the GPS L$_1$ band was shown in Figure 2.37(d). Notice that these measurements indicate antenna resonance only at L$_1$ whereas the return loss in the L$_2$ band for this antenna is poor.

2.3.1.3 Effects of Dielectric Core on a Self-Phased QHA

The cylindrical core used in the QHA is generally made either from foam ($\varepsilon_r \sim 1.07$) or a hollow, thin-walled dielectric tube made from a low-dielectric constant material such as Teflon or from periodic dielectric spacers. In some designs the
thin walled inner tube supporting the helical structure can be made from a high-temperature molded plastic. High-dielectric constant cores, such as ceramic cores, have been used for reducing the size of the QHA [59, 60]. The presence of the dielectric core alters the phase velocity of the RF current path along the helical windings and can affect the radiation pattern of the antenna [53]. Therefore, the pitch angle and the length of the helix need to be adjusted to compensate for these effects. In the uncompensated case, dielectric loading can cause the peak to occur at a lower-elevation angle and cause a significant loss in the peak gain. A reduction in length of the helix and the pitch angle effectively shortens the electrical path length and compensates for the phase error. Foo [53] has compensated for these dielectric loading effects by varying the pitch angle and reducing the overall length of the helix. Leisten [59] has used a cylindrical core made from very-high-permittivity ceramic material to design very-small-sized, lightweight QHAs that have been developed by Sarantel for use in GPS handsets. Liu and Chu [60] have recently designed an extremely compact triple-band, dielectrically loaded QHA intended also for handset applications. Using the high-dielectric core coupled with the antenna topology in both designs helps to insulate the antenna from interacting with the person carrying the handset and also with the surrounding environment. It does so by confining the near field to be closer to the antenna. This is an important type of antenna used in many GNSS handset applications and will be discussed in greater detail in Chapter 4. MOM techniques described in the previous section for analysis and design of air core or foam core QHAs cannot be used for designing QHAs with cores made from higher-dielectric constants. A transmission line method (TLM) such as Microstripes listed in Table 2.2 has been used by both Leisten [61] and Liu [60] for designing QHAs with ceramic cores. One major disadvantage in using ceramic cores for QHAs is the drastic reduction in bandwidth. Their bandwidth is so narrow that it allows them to receive only C/A code signals in the GPS L1 band but not the P(Y) code or M code signals in the GPS L1 or L2 bands, requiring either 20-MHz or 24-MHz bandwidths or even the L5 code, which requires a 20-MHz bandwidth.

### Table 2.4 Design Parameters for Three GPS Quadrafilar Helix Antennas Designed by Tranquilla and Best for Operation at 1.5754 GHz (L1 Band)

<table>
<thead>
<tr>
<th>Antenna Number</th>
<th>Axial Length ($L_{ax}$) Expressed in Wavelength $\lambda$</th>
<th>Radius of the QHA (b) in cm and in Wavelength $\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antenna 1</td>
<td>0.20</td>
<td>1.60 (or 0.084 $\lambda$)</td>
</tr>
<tr>
<td>Antenna 2</td>
<td>0.27</td>
<td>1.39 (or 0.072 $\lambda$)</td>
</tr>
<tr>
<td>Antenna 3</td>
<td>0.35</td>
<td>1.04 (or 0.054 $\lambda$)</td>
</tr>
</tbody>
</table>

From: [57].

2.3.1.4 Printed Self-Phased QHA

The four helical elements of the array of the QHA are made from either conducting wires (called WQHA for wire QHA) as shown earlier in Figure 2.36(a) or from metallic tapes using low-cost, printed circuit technology (PQHAs or printed QHAs). Extensive design work on PQHAs has been conducted by Sharaiha and coworkers [62]. The construction of a PQHA using low-cost, printed circuit tapes is illustrated...
in Figure 2.41(a) and (b) [55, 62]. The substrate used is a thin, flexible Kapton film (dielectric constant = 3.2 and about 50 µm in thickness) with electrodeposited copper cladding on its surface; four equally spaced, sloping-straight-line copper strips of width $w$ are etched on its surface. A strip of width $w$ is equivalent to a circular wire of diameter $4a$, where $a$ is the radius of a circular wire. The angle of each strip relative to the base line $\Delta$ is $\psi = \text{pitch angle of the individual helical antenna}$.

Figure 2.41  Construction of a printed self-phased QHA with a dielectric core.
elements of the QHA. The etched Kapton film containing the four lines are wrapped around a cylindrical core generally made from Rohacell foam whose circumference is equal to four times the interelement spacing $d$ between the helical elements. The axis of the cylinder core is normal to the baseline $\Delta$. The feeding arrangement for a self-phased QHA using tapes as the helical elements is shown in Figure 2.41(b).

### 2.3.1.5 Baluns for the QHA

Each of the two bifilars in the QHA operates as a balanced antenna when fed from a coaxial line through a balun. The function of the balun is to distribute equal currents but of opposite phase from the feed coaxial line to the two arms of the antenna. Bricker and Rickert [54] and Kumar [50] have proposed three different balun designs that can be used with the self-phased QHAs fed from the top. These are the folded balun, the split-sheath balun, and the infinite balun. The folded and split-sheath baluns are both fed from the center by coaxial cables mounted along the axis of the bifilar. A split-sheath balun is shown in Figures 2.36 and 2.37.

### 2.3.1.6 Printed QHA with Infinite Microstrip Balun Feeds

A new type of printed half-wavelength QHA consisting of microstrip fed helical elements with four opposite radial strips has been designed by Shumaker et al. [62]. Infinite microstrip baluns whose design is described in the paper is used to feed the helical elements. This design obviates the need for conventional folded or split-sheath baluns and reduces the physical size and weight of the antenna and makes it easy to mass produce. Experiments show that this antenna has a 3-dB bandwidth of more than 145° and a very impressive front-to-back ratio of 20 dB. This is a major improvement over QHAs with a conventional coaxial balun feeds described earlier.

### 2.3.2 Externally Phased QHA

The second type of QHA used in GNSS is the externally phased QHA shown in Figure 2.36(b) [50, 53, 54]. In this design the four antenna arms of the QHA are fed from the bottom of the antenna through an external broadband phasing network fed against a ground plane. Figure 2.36(b) shows a typical phasing network for the QHA consisting of 180° hybrid ring coupler and two 90° branch-line type couplers. The 180° hybrid ring coupler provides two outputs that are equal in amplitude but with relative phases of 0° and 180°, respectively. Each of these two outputs are in turn connected to the inputs of a 90° branch-line coupler to provide two outputs that have the same amplitude but with a relative phase difference of 90°. The feed network provides the four arms with equal amplitudes and with a relative phase of 0°, −90°, −180°, and −270°, resulting in RHCP end-fire radiation in the direction of winding. The external phasing method described above provides three advantages:

1. By eliminating the need for the narrowband balun, it allows a 50-ohm impedance match to the receiver over a much broader bandwidth
2. It allows the use of open-ended, equal-length helical elements and therefore allows the antenna designer to precisely tune the length of the helical ele-
ments by trimming the open end of each element or by varying the shape of each element.

3. It allows a dual-band design of the QHA by placing a resonant trap load at an appropriate position along the length of each open-ended helical element.

A dual-band QHA that operates in both the GPS L1 and L2 band has been designed on this principle [64]. Operation over dual-band designs is obtained by placing a resonant trap filter tuned to 1.5754 GHz (the center frequency of the GPS L1 band) at an appropriate position along the length of each open-ended helical element. The bandstop trap filter acts as a high-impedance “open” circuit in the GPS L1 band but as a virtual short circuit in the GPS L2 band, thus allowing the antenna currents to flow along the full length of helical element so as to achieve resonance in the lower L2 band.

2.3.2.1 Dual-Band and Triple-Band QHAs for GPS

A major disadvantage of the self-phased QHA as pointed out earlier was its narrow bandwidth indicated by Figure 2.37(c), which shows its measured return loss; hence, it is generally only designed to operate over only over a single GPS frequency band, generally the L1 frequency band of the GPS system. Other QHA designs have been proposed that allow dual-band operation over the GPS L1 and L2 bands. One proposed dual-band design [51, pp. 455–462; 57] uses two antennas, each operating at a different band, placed either in a piggyback configuration with the L1 antenna on top of the L2 antenna or by enclosing one antenna inside the other, with the L1 antenna inside the L2 antenna or by rotational offset with the L1 and L2 arms interleaved. Results of measurements performed on the first two proposed designs of dual-band QHAs have been provided by Tranquilla and Best [57]. However, neither of these proposed dual-band designs performs as well as the single-band version discussed earlier. Both designs show a reduction in gain in the upper hemisphere at both the bands, an increase in the backlobe level, and a poorer phase center performance at off-boresight angles when compared to the single-band design. In addition both designs appear physically clumsy and structurally suspect. There is no data provided on the third design of rotationally offset interleaved antennas. As mentioned in Section 2.3.1.3, a triple-band piggyback of ceramic-loaded QHAs for operation over the GPS L1, L2, and L5 bands, has been designed and built by Liu [60]; their design consists of a piggyback configuration of two separate antennas—a dielectric-loaded QHA (DQHA) at the top and a dielectric-loaded octa-filar helix (DOHA) at the bottom. The two antennas work in combination to operate in the GPS L1, L2, and L5 frequency bands. However, the bandwidth of these QHAs is extremely narrow because of ceramic loading—especially in the L1 band where only C/A code operation over a 2-MHz bandwidth seems possible from the data that has been provided.

A dual-band printed quadrifilar helix (PQHA) fed from the bottom with an external phasing network has been designed by Lamensdorf and Smolinski [64] by using trap filters placed at an appropriate distance along the length of the four helical elements of the PQHA. A picture of the dual-band PQHA is shown in Figure 2.42.
This antenna can provide a 24-MHz bandwidth in both the GPS L₁ and L₂ bands, which allows it to receive the P(Y) code and M code GPS signals. The antenna design is also physically more robust and provides good performance in both bands than the coaxial or piggyback dual-band versions mentioned earlier. The four helical arms are excited with sequential phase variation of 0°, −90°, −180°, and −270° in a counterclockwise manner to generate RHCP. The external phasing network is similar to that shown earlier in Figure 2.36(b). The trap filter consists of a parallel resonant LC circuit and is used as a switch to tune the antenna for resonance at two different resonant frequencies. The trap filter acts as an open circuit offering high impedance to current flow in the antenna at its parallel resonance frequency, which is selected to be 1.5754 GHz, the higher of the two selected frequencies for the QHA. The location of the trap filter in each helical arm is selected so as to provide resonance at 1.5754 GHz. At 1.2276 GHz, the lower of the two selected frequencies, the trap filter has low reactive impedance, allowing an almost unimpeded current flow in the antenna. The length of the antenna beyond the filter is compensated for the presence of the trap filter is tuned to achieve resonance at 1.2276 GHz, the center frequency of the L₂ band offering dual-band operation within a single integrated QHA. The four helical lines of the PQHA were etched on the thin copper-clad Mylar tape at the correct pitch angle. The tape was then wrapped around a cylinder made from Rohacell foam with an outer diameter of 0.36”. Discrete microstrip L and C components were used in building the resonant trap filter that was soldered across a small gap in each of the four arms of the PQHA. The PQHA was mounted on a circular ground plane with a diameter of 0.3 wavelengths at the lowest GPS frequency used in these measurements. The performance of the QHA was optimized using MOM techniques using the NEC-4 MOM code. Figure 2.43 shows the measured reflection loss S₁₁ in each of the two GPS bands. The bandwidth is 5% in the L₁ band and 2.2% in the L₂ band. Figure 2.44 shows the measured elevation plane pattern radiation pattern for elevation angles from zenith down to −120° over a 24-MHz bandwidth at frequencies of 1.563, 1.575, and 1.587 GHz in the L₁ band. Figure 2.45 shows the corresponding measured elevation plane radiation patterns GHz in the GPS L₂ band at frequencies

![Image of the antenna with labeled parts: Trap load for L1 band, Dielectric foam core, Bandpass trap load for L1 band.]

Figure 2.42 GPS dual-band externally phased QHA. (From: [64]. Courtesy of D. Lamensdorf.)
of 1.215, 1.227, and 1.239 GHz. Unfortunately, antenna patterns were not measured by the authors below $-120$ degrees; hence no information is available on the antenna backlobes. The phase of the antenna was also not measured so there is no information to determine its phase center stability.

### 2.3.3 A Summary of the Advantages and Disadvantages of QHAs for GNSS

Self-phasing QHAs, shown in Figures 2.36(a) and 2.37 are very popular for GNSS because of the many advantages that they provide. They are light in weight, small in
size, and inexpensive to build especially if photolithographic techniques are used in its construction, as illustrated in Figure 2.41. Their biggest advantage is that they do not need a ground plane, which makes them very compact. Yet another advantage provided by a self-phased QHA is its shaped-conical radiation patterns with very broad antenna beamwidth of 100° to 180° as shown in Figure 2.37(c). This makes the self-phased QHA ideally suited for receiving signals from multiple GPS satellites visible to the antenna in the upper hemisphere. The ability to acquire and track nearly all of the visible satellites enables the QHA to provide excellent PDOP. The pattern can be shaped in elevation through small changes to its geometrical parameters. Their size can also be reduced by using a dielectric core made from a high permittivity material such as a ceramic, although this may also reduce the bandwidth.

Self-phased QHAs also have several disadvantages. One principal disadvantage is that its cylindrical shape and longer length makes it unsuitable for avionics applications, where because of ARINC 743 size requirements a very-low-profile antenna of height no greater than 0.75" is required to reduce drag. Another major disadvantage is that it also needs a coaxial balun, which makes its bandwidth narrow as shown in Figure 2.37. Therefore, it is generally only able to cover a single GNSS frequency band, handicapping its use in the multiple frequency bands used in modernized GPS or Galileo. Piggybacked and coaxially enclosed versions of dual-band QHAs have been proposed but do not work very well [57]. Its phase center can also vary with elevation angle since its radiating surface is distributed over a longer length. There can also be some asymmetry in azimuth caused by the uneven lengths of the adjacent helical elements required to generate quadrature phasing. Another disadvantage is that when the QHA is used with baluns made from coaxial lines, it generally has a poor front-to-back ratio as shown in Figure 4.39(c); this can make the self-phased QHA more susceptible to multipath. More recently, however, a modified design using infinite baluns made from microstrip
lines has greatly improved the front-to-back ratios to 20 dB [63]. The use of small ground planes with a self-phased QHA is optional and can be used when there is a need to screen out multipath from ground reflections. The use of such a ground plane can, however, cause phase center excursions [60].

Externally phased dual-band QHAs provide greater bandwidths and when used with trap-loaded filters can perform over the GPS L₁ and L₂ frequency bands [64]. The disadvantage in this design is that it needs a small ground plane and also an external phasing network, making it heavier and bulkier than a self-phased QHA.

### 2.4 Hexafilar Slot Antenna for GPS

Antenna engineers from Garmin International have published a total of eight papers between 1996 and 1998 describing several different types of resonant cylindrical slot antennas that have been developed for GPS applications. These antennas are modifications of a conventional self-phased QHAs and use printed cylindrical slots in their construction. One of these new designs is a microstrip-fed, printed cylindrical slot antenna designed by Ho et al. [65]. References to several other types of resonant slot QHA designs developed by Garmin International are also given in Ho’s paper. This antenna consists of six quarter-wavelength resonant slots, each of which is rolled half-turn around a cylindrical laminate. The slot antennas maintain the required phase relation to each other needed to provide a RHCP pattern with near hemispheric coverage. Experimental results indicate a 3-dB beamwidth of more than 120° and a front-to-back ratio of more than 15 dB provide good resistance to multipath signals from the ground. Tests conducted on this antenna with the Garmin GPS 90™ receiver have confirmed its ability to track satellites at very-low-elevation angles for providing excellent PDOP. Figure 2.46(a) and (b) show the cross-sectional diagram of this antenna and its measured radiation pattern.

![Hexafilar Slot Antenna Diagram](image)

**Figure 2.46**  GPS hexafilar slot radiator: (a) cross section, and (b) measured radiation pattern (From [65]. ©1998 IEEE.)
2.5 Planar and Drooping Bow-Tie Turnstile Antennas for GNSS

Bow tie antennas are planar versions of the biconical antenna, which is well known for its broadband properties. RHCP needed for GNSS can be obtained from a pair of crossed bow-tie dipoles placed in the x-y plane over a ground plane and fed with equal amplitudes and in phase quadrature, with phases of 0° and −90°, respectively. The crossed dipoles are generally fed in shunt with either a split-sheath coaxial balun or a dual-folded balun; such a shunt connection will require two dipoles of different lengths with complex conjugate impedances needed to obtain the required 90° difference in phase. A physical arrangement of these crossed conjugate dipoles is called a turnstile configuration and it exploits the impedance properties of the dipole to obtain the required 90° phase difference. The impedance of the shorter dipole is capacitive, and its current will have a positive phase relative to its length at resonance. The longer dipole will have an inductive resonance and has a current with a negative phase relative to its condition at resonance. The relative phase quadrature needed to generate the CP can be obtained by optimizing the relative lengths of the crossed pair of dipoles. The dipoles need to be placed at an optimized height over a ground plane so that the LHCP signal they generate in the downward -z direction will upon reflection from the ground plane change polarization and add to the direct signal propagating upwards in the +z direction. Several factors such as optimizing the lengths of the dipoles, the height above the ground plane, and the finite bandwidth of the balun will place restrictions on taking full advantage of the maximum bandwidth capability provided by the bow-tie antenna geometry. Pattern coverage in the elevation plane can also be “tailored” by drooping the bow ties down towards the horizon to provide the gain needed for receiving signals from the low-elevation GNSS satellites. Planar as well as drooping crossed bow-tie dipole antennas on top of both planar and conical ground planes have been investigated by Kaswarra et al. [66] and Makarov [67] for modernized GPS and Galileo at frequencies ranging from 1165 to 1300 MHz. This covers the L5 and L2 bands of modernized GPS as well as the neighboring Galileo bands from 1250 to 1300 MHz. The geometry of these antennas is shown in Figure 2.47. The planar and drooping dipole versions of their antenna placed over conical ground planes are shown in Figure 2.48. Their investigation indicates that the best polarization bandwidth is obtained by a drooping bow-tie turnstile placed over a planar ground plane. The dimensions of the antenna were optimized using the HFSS finite element method (FEM) solver for an elevation angle θ = 60° (elevation of 30°) and for azimuth angles φ = 0°, 60°, 90°, and 135°. A flare angle of 75° for the bow tie and a drooping angle of 30° provided the best performance with the antenna placed on top of a planar ground plane at a height \( h/\lambda_0 = 0.26 \) where \( h \) is the height above the planar ground plane and \( \lambda_0 \) is the wavelength corresponding to a frequency of 1.2 GHz, the center frequency of the frequency band that was needed for these applications band. The diameter of the planar ground plane was set to 0.8 \( \lambda_0 \) and for the conical ground plane the height to diameter ratio was set to 1.4. Optimum polarization performance was obtained when \( l_y = 0.25 \lambda_0 \) and \( l_x = 0.13 \lambda_0 \), where \( l_y \) and \( l_x \) are the half-lengths of the dipole. The measured return loss for the antenna is shown in Figure 2.48. The use of cavity-backed planar microstrip bow-tie dipoles for Galileo
has been mentioned by Granger et al. [68]; however, no detailed description has been provided in the paper of either the design or the measured performance of this antenna.
2.5.1 End-Fire Array of Turnstile Antennas for Multipath Limitation

An end-fire array of three identical turnstile dipole antennas has been used by for limiting multipath effects on a GPS receiver by Counselman [69]. This array directs the radiation towards zenith while suppressing backlobes to reduce multipath signals reflected off the ground. It obviates the need for a ground plane to suppress the antenna backlobes. The array is quite complex and consists of three antenna elements spaced one-third of a wavelength apart at the center frequency of the respective GPS bands. Each array element consists of separate turnstile antenna elements covering the L1 and L2 bands with 90° hybrids to generate RHCP. The turnstile elements are crossed horizontal dipoles; they are linear in shape and do not have bow-tie shapes to make them broadband as in the turnstile antennas considered in the previous section. The L1 and L2 arrays are connected to a GPS receiver through a L1-L2 duplexer. The array was tested at both the GPS L1 and L2 frequency bands and the multipath reject performance was compared against that of a planar metal ground plane. The array is approximately 0.1m in diameter, is 0.4m in height, and resembles a vertical post. Tests indicate that the three-element array was able to reject the multipath about 5 dB better than a 0.5m diameter metal ground plane. When the number of elements in the array was increased to five elements the improvement in multipath rejection was 5 dB better than obtained by a 0.9m diameter metal ground plane.

2.6 Directional GNSS Antennas

The design of three types of directional antennas used for GNSS applications: helical, conical spiral, and parabolic reflector antennas are described in this book and shown in Figure 2.1. Two of these antennas—helical and parabolic reflector—will be discussed in this chapter. Conical spiral antenna design along with other types of multiband spiral antennas are discussed in Chapter 4.

2.6.1 Helical Antennas

Two types of circularly polarized helical antennas are used in GNSS:

1. Cylindrical axial mode helix antenna with wide bandwidth and relatively high gain in the boresight direction [66] is shown in Figure 2.49(a);
2. Hemispherical low-profile helix antenna with a broad CP beam for providing coverage over a larger part of the upper hemisphere similar to a FRPA [70, 71], which has lower gain than a helix, is shown in Figure 2.49(b).

The hemispherical helix antenna, however, has an impedance bandwidth of only 6% as compared to the axial mode helix, which can be optimized to have a bandwidth of 34% to cover the entire GNSS spectrum if necessary. The hemispherical helix can only cover a single GNSS band and has normally been designed to cover just the GPS L1 band. It has been used as an element in a four-element GPS beamforming antenna array for providing some mitigation of multipath and interference [71].
2.6 Directional GNSS Antennas

2.6.1.1 Cylindrical Axial Mode Helix

The cylindrical, monofilar helical antenna, made from either a single wire or metalized tape, can provide a CP pattern with a relatively high gain over a fairly wide bandwidth. Figure 2.49(a) shows a picture of an axial mode helix for applications in the modernized GPS covering the L5, L2, and L1 frequency bands [44]. It has been used as a stand-alone transmitting antenna or more frequently as an element in a transmitting antenna array in GNSS satellites. The narrow beamwidth and larger physical size and weight of the axial mode helix antenna generally precludes it from being used as a receiving antenna that can be integrated into GNSS receivers but it is popular as a stand-alone transmitting antenna in many tests and laboratory measurements. An interesting recent application for the axial mode helix has been as a transmit antenna in indoor navigational systems using GPS-like pseudolites (PLs) that are capable of providing centimeter accuracy [4]. A microstrip patch antenna cannot be used in such applications because its beam is too wide and can create serious multipath from interaction with the surrounding environment. In contrast, the beamwidth of the axial mode helix can be tailored to the desired narrower beamwidth by adjusting its length, diameter, and other parameters; hence, it is the preferred antenna of choice for such measurements.

2.6.1.1.1 Helical Antenna Arrays Used in GPS Satellites

Helical antennas with either cylindrical or conical cup ground planes are also very popular as transmitting elements in antenna arrays used in several GPS navigational satellites. There are four different types of satellite antennas: (1) Block I, (2) Block II, Block IIA, (3) Block IIR-A, and (4) Block IIR-B and Block IIR-M [1–3]. Figure 2.50 shows helical antennas used on the GPS IIR-M satellites built by Lockheed Martin Co. [1]. Figure 2.51 shows the helix antennas used in the Block IIA satellites by Rockwell Collins [2]. A summary of the antenna arrays of axial mode helical elements used as transmitting antennas in the GPS satellites—Block I, Block IIA, and Block IIR—are described by Aparicio et al. [3]. The bandwidths of the elements in these earlier satellites range from 1200 to 1600 MHz and hence do not include...
the L5 signal. Each array contains twelve RHCP helical elements arranged in two
concentric circles, with four elements in the inner circle and the remaining eight
elements in the outer circle [72]. The earlier Block I satellites used a cylindrical cup
ground plane but the later systems, the Block IIA and IIR, have used conical cup
ground planes. The PCV of these satellite antenna arrays have also been studied by
Schmid et al. [72] from the vast amount of data that they have collected. The PCV
for the Block IIR-A, Block I and Block II and IIA satellite antennas was between ±5
mm, whereas for the Block IIR-B the PCV was almost double this value of nearly
±10 mm.

2.6.1.1.2 Design of Cylindrical Axial Mode Helix Antennas

The axial mode helix is a traveling wave antenna and consists of two parts: the
director, which is the main helical portion, and the exciter/ground plane, which is
the feed portion for the antenna. The helix can be made either from a single wire, a
thin tube, or a flat strip conductor on a thin dielectric tape; these are wound around
a supporting dielectric cylinder. Tapes are commonly used in satellite antennas, as
shown in Figures 2.50 and 2.51, because of their lighter weight. The polarization
is determined by the direction of the windings; thus a right-hand screw winding of
the helix wire produces the required right-hand polarization needed for GNSS. The
direction of the right screw winding for the helical wire is determined by pointing
the thumb of the right hand along the desired direction of propagation and by
wrapping the four other fingers in the direction of the wire winding. The geometry
of the axial mode helix antenna with either a cylindrical or conical cup ground
plane is shown in Figure 2.52. The performance of the helical portion is controlled
by the various geometrical parameters of the helix antenna listed in Table 2.4.
The early “classic” design models for this antenna [74] provided high gain but
over a relatively narrower bandwidth. They assumed that the radius of the helical
wire would have an insignificant effect on antenna performance; this assumption

Figure 2.50  GPS IIR-M satellite antenna assembly. (Printed with permission from Lockheed Martin
Co.)
2.6 Directional GNSS Antennas

turned out to be erroneous in achieving the maximum possible bandwidth. New techniques for designing uniform diameter helical antennas have been proposed by Djordjevic [75]; these take into account the influence of the radius of the wire for achieving the maximum possible bandwidth and they provide a lower peak gain but over a much wider bandwidth than the classical designs. The trade-off between gain and bandwidth in these two designs are explained in Figure 11 of the paper by Djordjevic [75]. Using these new wideband design techniques the performance of the helical antenna can be optimized to provide good gain over the entire GNSS spectrum covering the GPS, Galileo, and GLONASS systems with frequencies ranging from 1.145 GHz (the lower-band edge of Galileo E5 band) to 1.615 GHz (the upper-band edge of the GLONASS L1 band) for a total bandwidth of 34%. Several other parameters of the helix can also be varied to achieve even further improvements in the pattern and gain bandwidth. For example, the diameter of the cylinder can be tapered or made nonuniform, the pitch of the helical winding can be changed along the axis, or the diameter of the helical wire or the width of the helical flat strip conductor can be varied [75]. The return loss of a uniform diameter helix as a function of frequency can be improved simply by adding two additional tapered turns at the free end of the helix as in the satellite antennas for the Block IIR-M satellites shown in Figure 2.50.

The exciter portion of the helical antenna consists of a ground plane with either a cylindrical or a truncated conical cup at the base of the helix as shown in Figure 2.52. These improve the gain, front-to-back ratio and sidelobe level, thereby
reducing the susceptibility of the antenna to multipath. Inside this exciter portion the helix wire is connected to a coaxial connector mounted on the ground plane. The helical antenna has an input resistance \( R_{in} = 141 \frac{C}{\lambda} \) ohms where \( C \) = the circumference and \( \lambda \) = the wavelength at the midband frequency. A wideband impedance matching section is needed to match the helix to standard 50-ohm coaxial connector. This can be designed by using several methods such as by bringing the last quarter turn of the helix wire or tape parallel to the ground plane in a gradual manner and inserting a microstrip matching transformer between the coaxial feed and the beginning of the helix wire \[76\] or alternatively by using a vertical-profiled metal plate at this junction \[75\]. The optimum diameter \( D \) of a cylindrical cup used with the ground plane is \( D = 1 \lambda \) and the height of cup is \( h = 0.25 \lambda \). The cup can increase the antenna gain by 1.4 dB compared to a planar ground plane. The optimal dimensions for the truncated conical cup are as follows: the diameter at the base of the cone is \( D_1 = 0.75 \lambda \); the diameter at the apex of the cone \( D_2 = 2.5 \lambda \); the height \( h = 0.5 \lambda \). The cone suppresses the sidelobes near the horizon and also the back radiation but does not appreciably increase the gain in the boresight direction. Near the exciter feed region the current on the helix attenuates smoothly to a minimum, while the current over the remaining length of the helix is relatively uniform, allowing the current on the helix to be approximated as being relatively uniform while computing its radiation pattern.

2.6.1.2 Design and Measured Performance of an Axial Mode Helical Antenna for Modernized GPS

A picture of a helical antenna with a cylindrical cup ground plane built by the MITRE Corporation \[44\] is shown in Figure 2.49(a). The design parameters for this antenna are listed in Table 2.5. This antenna is capable of operating in the L5, L2, and L1 bands of the modernized GPS system and has a 31% bandwidth; it can also...
be used in the neighboring Galileo E5 and E6 bands from 1145 to 1.300 GHz. Its use thus far has been as a transmitting antenna to evaluate the fidelity of M code reception of GPS receiving antennas through correlation measurements in the L1 and L2 bands [44]. The radiation patterns and gain of this antenna measured at 1.176, 1.227, and 1.575 GHz, and the center band frequencies of the L5, L2, and L1 bands are shown in Figure 2.53. The measured pattern shows good performance at the two lower frequencies but there is some degradation in gain performance and broadening of beamwidth at 1.5754 GHz. The antenna was designed using formulas from classical narrowband design mentioned earlier without accounting for the effects of the diameter of the helical wire on the bandwidth. It would require only a small variation in its design to increase its bandwidth to the 34% needed to cover the entire GNSS spectrum. The design methods recommended by Djordjevic [75] can be used as discussed below for making such an improvement. The classical design recommends the following guidelines for varying the antenna parameters listed in Table 2.5 for achieving optimum performance [74]:

- Normalized circumference $C/\lambda_C$ should be between 0.71 < $C/\lambda_C$ < 1.20, where $C$ is the circumference and $\lambda_C$ is the wavelength at the midband frequency;
- Pitch angle $\alpha$ be between $12^\circ < \alpha < 15^\circ$, where $\alpha = \arctan (p/C)$;
- Normalized pitch or spacing between turns $p/\lambda_C$ should be between 0.15 < $p/\lambda_C$ < 0.3.

The number of turns $N$ should be between $4 < N < 15$. These recommendations are based on the assumption that the antenna performance is independent of the
radius of the wire. Other parameters needed for completing the design and building the antenna are the total length of the antenna \( L = Np \) and the total length of the antenna wire \( L_w = N \sqrt{p^2 + C^2} \) where \( \sqrt{p^2 + C^2} \) is the length of the wire between each turn. The broadband design [75] takes into account the wire radius as well as the total length “\( L \)” of the helix. (The optimum values for these same parameters can be obtained from Figures 4 and 6 and the desired bandwidth, axial ratio, and maximum gain can be determined from Figures 3, 5, and 7, respectively, provided by Djordjevic [75].) The antenna shown in Figure 2.49(a) has a normalized wire radius \( r/C = 0.0059 \), where \( r \) = radius of the helical wire. Increasing the diameter of this helix to 1.574" from its current radius of 1.4" and the reducing the pitch angle to 13° from its current value of 14° would increase the bandwidth to the 34% needed to cover the GNSS spectrum. The length \( L \) can be adjusted to achieve a higher gain.

Empirical formulas for calculating the maximum gain, the optimum radius of the helix, the half-power beamwidth (HPBW), and the optimum axial ratio are also available. Emerson [73] has proposed a simple empirical expression for the maximum gain \( G_{\text{max}} \) (dBic) based on numerical modeling of the helix as a function of the total length of the helix \( L \) normalized to the wavelength \( \lambda_{c} \); this formula gives results that compare well with measurements:

\[
G_{\text{max}} \text{ (dBic)} = 10.25 + 1.22 \left( \frac{L}{\lambda_{c}} \right) - 0.0726 \left( \frac{L}{\lambda_{c}} \right)^2
\] (2.57)

The normalized optimum radius of the helix for maximum gain can be obtained from the wideband theory of Djordjevic [75].
\[ \frac{a}{\lambda_c} = 0.2025 - 0.0079 \left( \frac{L}{\lambda_c} \right) - 0.0726 \left( \frac{L}{\lambda_c} \right)^2 \]  

(2.58)

The empirical equation for the half-power beamwidth HPBW is [74; pp. 12-11]

\[
\text{HPBW} = \frac{61.5 \left( \frac{2N}{N + 5} \right)^{0.6}}{\left( \frac{2\pi a}{\lambda} \right)^{0.7}} \left( \frac{\tan \alpha}{\tan 12.5^\circ} \right)^{\frac{N}{4}} \text{ (degrees)}
\]  

(2.59)

The polarization of the normal mode helix is in general elliptical with an axial ratio given by

\[ AR = \left| \frac{E_\theta}{E_\phi} \right| = \frac{2p\lambda_c}{C^2} \]  

(2.60)

The antenna will be circularly polarized if \( AR = 1 \) is satisfied. This condition is satisfied when the circumference of the helix and the pitch or spacing between the turns are related as \( C = \sqrt{2p\lambda_c} \), which in turn be used to find the pitch angle of the helix as

\[ \tan(\alpha) = \frac{p}{C} = \sqrt{\frac{p}{2\pi \lambda_c}} \]  

(2.61)

The far-field radiation pattern of the axial mode helical antenna can be determined by considering the \( N \) turn helix as an array of \( N \) identical elements with an interelement spacing equal to \( p \). The normalized field pattern is then obtained by multiplying the pattern of one turn of the helix by the array factor. The result is

\[ E(\theta) = \sin \left( \frac{\pi}{2N} \right) \cos(\theta) \frac{\sin \left( \frac{N\psi}{2} \right)}{\sin \left( \frac{\psi}{2} \right)} \]  

(2.62)

In the above equation \( \cos(\theta) \) is the element pattern and \( \frac{\sin \left( \frac{N\psi}{2} \right)}{\sin \left( \frac{\psi}{2} \right)} \) is the array factor of a uniform array of \( N \) equally spaced elements \( \psi = \frac{2\pi}{\lambda_c} p \cos \theta + \alpha \). Here \( \alpha \) is the phase shift between successive elements and is given as \( \alpha = -2\pi - \frac{\pi}{N} \).

### 2.6.2 Helibowl Multipath Limiting Antenna

The Helibowl is a bowl-shaped reflector antenna that is excited by a helical antenna element located at the center of the bowl [77]. The Helibowl was originally
developed by Don Spitmesser at the Jet Propulsion Laboratory in Pasadena, California [77] by modifying a low sidelobe “helicone” antenna concept first proposed by K. Carver [78]. It provides high and uniform gain over the desired range of elevation angles above a certain masking angle followed by a sharp cutoff in gain below the masking angle. The Helibowl is used in conjunction with a phased array to track low-elevation satellites and has been extensively investigated at Ohio University [79, 80]. Because of its low sidelobe and backlobe levels and its ability to limit multipath, this antenna is used frequently in DGPS ground reference stations.

2.6.3 Hemispherical Helix Antenna and Its Use in a GPS Beamforming Array

The axial mode helix described in the previous section provides a good axial ratio but over a narrow beamwidth because of its relatively high gain. Its radiation pattern is not that of a “classic” FRPA and will not be optimum for receiving signals simultaneously from four or more GPS satellites spread over a wide range of viewing angles in the upper hemisphere. A five-turn, printed hemispherical helical antenna for GPS applications has been designed by Zhang and Hui [70] and is shown in Figure 2.49(b); it provides a much broader, CP antenna pattern for acquisition of multiple GNSS satellites. It has a low profile that allows a stable mechanical installation above a ground plane and an impedance bandwidth of 6% with a return loss of greater than 20 dB. It also provides a 3-dB axial ratio bandwidth of between 6% to 7% and has a gain of 9 dB at its center frequency. The simulated patterns show good CP over a ±45° range of elevation angles around zenith, but also show some asymmetry with variation of azimuth angle since the patterns measured at orthogonal azimuth angles of 0° and 90° are different at certain elevations. An input impedance matching network for the GPS L1 band is provided by a three-element matching network also shown in this figure; the return loss it provides varies from −10.7 to −27.7 dB across the band.

The hemispherical helix (a wire version) has been used as an antenna element in a four-element GPS beamforming antenna array to mitigate multipath and interference effects [71].

Figure 2.54(a) shows a picture of the four-element helical antenna array. The radius of each hemispherical antenna element in the array is 37.5 mm and the wire radius is 0.5 mm. The height of the array is not given. Measured beam patterns of the four antenna lobes generated by the array are shown in Figure 2.54(b). The elements are arranged in a square lattice separated at a distance of 0.7 wavelengths at 1.5754 GHz along both the x and y axes. An inexpensive, analog Butler Matrix feeding network made from four branch-line quadrature hybrids is used in the array. The array can generate four spatial beams along two orthogonal directions; the beam directions are at azimuth angles of ϕ = 45°, 135°, 225°, and 315°, respectively and at an elevation angle θ≈25°. The half-power beamwidth of each beam is ~44° and the gain is greater than 17 dB; the axial ratios of the beams are between 1.5 to 2.2 dB.

2.6.4 High-Gain Reflector Antennas for Monitoring GNSS Signals

Large reflector antennas providing very high gain are increasingly being used both in the United States and in Europe for monitoring the quality and spectral purity
of signals transmitted by recently launched satellites. These antennas also are being used for identifying signal interference and for characterizing potential signal anomalies and errors caused by the satellite hardware. The received signal power spectral density of GNSS spread spectrum signals is approximately 20 dB below the thermal noise floor. The baseband signal bandwidth is intentionally spread over a wide bandwidth and appears as noise; the original signal is recovered through a despreading process in the receiver that can provide between 10 to 60 dB of processing gain. Before this despreading the chips of the C/A and P(Y) codes are not discernible. The high gain of these large reflectors when coupled with new “dithered sampling frequency signal processing techniques” provides a positive signal-to-noise ratio (SNR) and allows both the individual chips and the navigation bits to be demodulated without performing the despreading procedure.

These types of monitoring investigations started in the early 1990s and were first used to diagnose an anomaly observed in satellite vehicle number (SVN) 19 [5]. It has since been used to monitor signals from recently launched satellites such as the Galileo satellites (GIOVE-A & B) [7] and the Compass M1 satellite [6, 81]. The largest of the reflector antennas used for this purpose is the 110m diameter Robert C Byrd Telescope antenna at Green Bank, West Virginia, which provides a gain at L band of about 70 dB [5]. Several of these large reflector antennas currently being used for these monitoring investigations are described below.

The GNSS Monitor Station (SGMS) at Stanford University, which uses two reflector antennas, [6, 81] has been very active in this field for the past few years. The smaller of these two is a front-fed parabolic reflector antenna 1.8m in diameter. Its estimated gain, assuming an aperture efficiency of 50%, is 23.3 dBi at 1.1 GHz and 26.75 dBi at 1.6 GHz. Since the maximum gain of a typical FRPA is around 1 to 3 dBi, this reflector antenna can provide a boost in gain (or in SNR) of nearly
20 to 23 dB across the GNSS band over a FRPA. The estimated HPWB is 8.85° at 1.1 GHz and around 6.08° at 1.6 GHz. The second dish at SGMS is even larger: a 45.7m (150 feet) reflector antenna called the SRI Stanford “Big” dish [81]. This antenna has a 50 dB of gain and a beamwidth of 0.25 degrees with an efficiency of 35% and it operates from 1.2 to 1.6 GHz.

Figure 2.1 shows a Cassegrain dual-reflector antenna, with a 30m diameter primary parabolic reflector and a 4m diameter hyperbolic subreflector is being used by the German Aerospace Center (DLR) in Weilheim, Germany [7]. Its gain is estimated to be 50 dBi with a HPWB of 0.5° [7]. The antenna has a broadband feed covering 1.1 to 1.6 GHz and a position accuracy of 0.001° and has been used for analysis of signals from the Galileo satellites. A parabolic reflector antenna 5.1m in diameter is currently being acquired by Alcatel Alexia Space Italia (AAS-I) for their NAVCOM laboratory [8]. This antenna is also a front-fed parabolic reflector with a diameter of 5.1m. Its gain is estimated by AAS-I to be 32 dBi at 1.1 GHz with a half-power beamwidth of 4°; its gain at 1.6 GHz is 36.9 dBi with a half-power beamwidth of 2.75°. Its G/T is 11.5 dB/K pointed at zenith and is 5.7 dB/K pointed at the horizon. Other notable examples of reflector used for GNSS monitoring antennas are a 2.4m antenna at the Toulouse Space Center (CST) in France and two dishes of 12 and 7m at Monitoring Earth Station in Leeheim, Germany [81]. Several other large radio telescopes in Europe are also being used in GNSS monitoring.

Most of these antennas are fed by broad L band feeds, allowing them to monitor signals transmitted by all current GNSS satellites including GPS, Galileo, GLONASS, and Compass. These reflector antennas can be one of two types: either a front-fed parabolic reflector or a Cassegrain dual reflector consisting of primary reflector and a subreflector. A detailed description of the design of these single- and dual-reflector (Cassegrain) antenna systems is beyond the scope of this book but is well described in two recent books [82, 83]. The last of these references [83] is devoted exclusively to a description of very large reflector antennas similar to those used in GNSS monitoring stations.

The gain $G$ of these parabolic reflector antennas is determined mainly by the size of the aperture of the primary reflector and the illumination amplitude taper across the aperture to reduce the sidelobe level. The gain $G$ can be calculated from:

$$G = \frac{\varepsilon \pi^2 D^2}{\lambda^2}$$

where

$D = \text{diameter of the reflector}$

$\lambda = \text{wavelength}$

$\varepsilon = \text{aperture efficiency, which is a function of several factors such as spillover loss from the feed, amplitude taper across the aperture to reduce sidelobe level, blockage from struts supporting the feed antenna, random surface errors, and so forth.}$
The efficiency can vary from 35% for a low-efficiency feed to 70% for better reflector designs. The HPBW in degrees for a uniform aperture taper across the primary reflector can be calculated from:

$$HPBW = 58.44 \left( \frac{\lambda}{D} \right)$$

For the 1.8m Stanford reflector the HPBW is 8.85° at 1.1 GHz and 6.08° at 1.6 GHz. If for the purpose of reducing the level of the antenna sidelobes the taper across the aperture of the reflector were to be made quadratic instead of uniform, it would cause a decrease the gain and a corresponding increase in the beamwidth. Thus at a frequency of 1.6 GHz if the main reflector has an edge illumination of −10 dB relative to the center to decrease its sidelobe level its beamwidth can increase to 6.8° but its sidelobe level will decrease from −17.6 dB for a uniform illumination to −22.3 dB.

2.7 Beamforming Antenna Arrays

A beamsteering antenna array can steer four or more multiple high-gain beams towards GNSS satellites to receive only signals of interest (SOI) while simultaneously mitigating signals not of interest (SNOI) from either multipath or interference; it is able to differentiate between these two types of signals by using their spatial separation. Digital beamforming antenna arrays are frequently used in wireless communications for improving signal quality.

The beamsteering principles of this type of antenna array is best illustrated by first considering a 3-D array of N antenna elements whose geometry is shown in Figure 2.55.

The azimuth angle is measured counter clockwise from the X axis and an angle θ measured from the Z axis; θ = 90° represents the horizon. The location of the element n in the array is \( \mathbf{r}_n = x_n \hat{x} + y_n \hat{y} + z_n \hat{z} \) where \((x_n, y_n, z_n)\) can be considered as the coordinates of the phase center of the nth antenna element in the array. One of the centrally located antenna elements in this array, \( n = 1 \), is designated as the “reference” element and is located at the origin. The location of this reference element is \( \mathbf{r}_1 = (0) \mathbf{x} + (0) \mathbf{y} + (0) \mathbf{z} \) is the wave number of a signal—the SOI—arriving at the array from the satellite at azimuth \( \phi_s \) and at an angle \( \theta_s \) measured from zenith and is represented by

$$\mathbf{k}_s = \beta (\cos \phi_s \sin \theta_s \hat{x} + \sin \phi_s \sin \theta_s \hat{y} + \cos \theta_s \hat{z})$$

here \( \beta = \frac{2\pi}{\lambda} = \frac{2\pi f_0}{c} \) and \( c \) is the velocity of light = 3 \times 10^8 \text{ meters per second}. The difference in phase between the satellite signal incident on an element \( n \) in the array and the reference element 1 at the origin is

$$\Delta \psi_n = \beta \cdot \Delta d_n = \beta \mathbf{k}_s \cdot (\mathbf{r}_n - \mathbf{r}_1) = \beta (x_n \cos \phi_s \sin \theta_s + y_n \sin \phi_s \sin \theta_s + z_n \cos \theta_s)$$
The signals received by all $N$ elements in the array compared to the signal in the reference element 1 can then be represented in terms of a steering vector $S(\theta_S, \phi_S)$, which describes the phase shifts of the array elements needed to steer the main beam of the antenna array in the direction $(\theta_S, \phi_S)$ of the SOI from the satellite.

$$S(\theta_S, \phi_S)^T \left[ 1, s_1(\theta_S, \phi_S), \ldots, s_N(\theta_S, \phi_S) \right]$$

$$s_n(\theta, \varphi) = e^{-i\Delta \psi_n}$$

It is also assumed that incident signal and each element in the array are copolarized so that there is no polarization loss on reception. The modulation of the incident satellite signal will be represented by its baseband complex envelope $a(t)$. The signal $s(t)$ at the output terminal of the array manifold will then be

$$s(t) = a(t) \sum_{n=1}^{N} g_{na}(\theta_S, \varphi_S, f_0)e^{-i\Delta \psi_n}$$

$g_{na}(\theta_S, \varphi_S, f_0)$ is the gain of the $n$th array element along a direction specified by $(\theta_S, \varphi_S)$. In an actual array the gain $g_{na}(\theta_S, \varphi_S, f_0)$ may not be the same for all elements in the array. It is necessary to accurately account for the differences in the phase and amplitude weight contributions of the individual antenna elements through calibration if optimum results are to be obtained from beam forming.
discussed in this section. However, to simplify the analysis we will make a number of initial assumptions that rarely, if ever, occur in a real antenna array. We will ignore mutual coupling effects between the antenna elements in the array and also the diffraction effects from the ground plane or the airborne platform on which the array is mounted. We will assume that all antenna elements in the array have identical radiation patterns. Furthermore, we will also assume that all array elements are RHCP with the same axial ratio. The gain of each of the $N$ antenna elements for the principal RHCP under these simplifying assumptions will be represented as $g_d(\theta_S, \varphi_S, f_0)$. Hence,

$$s(t) = a(t) \sum_{n=1}^{N} g_{na}(\theta_S, \varphi_S, f_0) e^{-j\psi_n} = a(t) g_d(\theta_S, \varphi_S, f_0) f(\theta_S, \varphi_S)$$

The term $f(\theta_S, \varphi_S)$ is called the array factor; it determines the ratio of the received signal available at the array output to the signal measured at the central reference element of the array.

To steer the main beam pattern of the array, each antenna element in the array is assigned complex filter weight $w_n$ that has both a magnitude and a phase component associated with it. The magnitude of weighting on the array elements is varied only if it is desired to reduce the sidelobes of the array pattern, otherwise it can be the same for all elements. The weight vector is represented by:

$$w^T = [w_1, w_2, \ldots, w_n]$$

In the vector notation the weighted sum of the total input signal received by the array is then represented by

$$u(t) = w^Tu$$

where

$$v^T = [v_1, v_2, \ldots, v_n]$$

The output of the array manifold from the signal received from the GNSS satellite at $(\theta_S, \varphi_S)$ is

$$u_s = w^Tv$$

where

$$s^T = [s_1, s_2, \ldots, s_N]$$

The received power at the array output

$$P_s(\theta_S, \varphi_S) = E\left| \sum_{n=1}^{N} w_n a(t) g_d(\theta_S, \varphi_S, f_0) e^{-j\psi_n} \right|^2$$
where $E[\cdot]$ is the expectation. By adjusting the complex weights $w_n$ on each of the $N$ elements it is possible to do “beamforming,” which is directing the main beam of the array manifold in the direction $(\theta_S, \phi_S)$ of the incoming signal to maximize the signal received by the antenna array from any visible GNSS satellite. In a real-world antenna array, the complex weights $w_n$ on each array element can also be optimized if necessary to compensate for the presence of mutual coupling or any change in the gain or polarization state of the antenna element in the array through careful precalibration.

If the positions of the GNSS satellites relative to the antenna elements of the array are known it would be possible to steer the main beam of the array in these directions. This also requires that the position and the orientation of the array is also known. The adaptive beamforming antenna is therefore designed as an integrated INS/GNSS unit. The outputs provided by the INS unit are the position and attitude of the array relative to the earth, which is generally obtained from the inertial measurement unit (IMU). Using this information the beamsteering unit in the array can calculate the directions of the satellites relative to the array and select the appropriate complex weights for each element. A wide range of different beamforming algorithms have been proposed in literature; these are based on using a priori information to discriminate between the SOIs, which are the satellite signals, and the SNOIs, and also the correlation properties of the GNSS signals.

Figure 2.56 shows a schematic diagram of a generic beamforming GNSS antenna array.

The beamforming array of the type described above provides several advantages:
• Under ideal conditions the array can improve the satellite signal by a factor of $\log(N)$, where $N$ is the total number of antenna elements in the array.

• Since each beamformer is directed towards a specific satellite, it avoids phase discontinuities in the received GNSS signal that can occur in a single beamformer array thereby allowing more accurate carrier and phase differential tracking techniques.

• Beamforming antenna arrays have also been investigated recently by several authors as a method for reducing multipath effects in GNSS systems [10]. They use the spatial separation between the satellite signals and the multipath signals to vary the antenna array pattern so as to minimize multipath signals while enhancing the satellite signals.

Despite the many advantages that they provide, beamforming antenna array also has some significant disadvantages:

• This type of antenna array can become complex, large, and expensive since it needs a separate beamformer for each GNSS satellite and a minimum of four beamformers are needed for the array to function.

• The spatial environment of a highly dynamic vehicle can change rapidly resulting in large variations in attitude and directions of SOI. This requires that the complex weights of the array elements need to be constantly calculated and updated; but every change in the weight vector can result in an abrupt phase shift to the signal delivered to the receiver, which could create problems for the tracking loop.

• Another disadvantage is that each beamformer output needs to be passed independently to the GNSS receiver. This means that it cannot be used with most current GNSS receivers without making significant design changes although a software-defined radio is a good candidate for accomplishing this task.

Figure 2.57  Two high-gain GPS beamforming antenna arrays (HAGR) built by NAVSYS. (a) Beamforming array of 100 elements, and (b) beamforming array of 16 elements. (Figures provided courtesy of A. Brown of NAVSYS Corp.)
A beamforming array built by the NAVSYS Corporation has recorded signal improvements of between 10 to 20 dB depending on the number of elements in the array [9, 10]; a picture of a very large and a smaller beamforming GNSS antenna arrays built by NAVSYS is shown in Figure 2.57. Various computer and hardware simulations have been performed by Konovaltsev et al. [84] to demonstrate the efficacy of various beamforming algorithms. The beamforming is conducted in the digital domain and requires downconversion of the L band GNSS signals in the receiver front end.

References

[1] Lockheed Martin Space Systems Co. 80215-8500, Littleton, CO. (Permission for pictures shown in the book was granted by Stephen O. Tatum.)


